



Political economy of Ramsey taxation

Daron Acemoglu^a, Mikhail Golosov^b, Aleh Tsyvinski^{b,*}

^a MIT, United States

^b Yale and NES, United States

ARTICLE INFO

Article history:

Received 23 August 2009
Received in revised form 29 August 2010
Accepted 26 October 2010
Available online 3 December 2010

JEL classification:

H11
H21
E61
P16

Keywords:

Capital taxation
Fiscal policy
Political economy

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 politician is as patient as 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.

© 2010 Published by Elsevier B.V.

1. Introduction

Atkeson et al. (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 et al., 1997), models where capital-holders are distinct from workers (Judd, 1985), and certain overlapping generations models (Atkeson et al., 1999; Garriga, 2001; Erosa and Gervais, 2002). Similar results hold in stochastic versions of the neoclassical growth model (e.g., Zhu, 1992; Chari et al., 1994).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. To start with, 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 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. Despite this first-order effect, we also 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.

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

* Corresponding author.

E-mail address: a.tsyvinski@yale.edu (A. Tsyvinski).

¹ A 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 et al., 2003; Kocherlakota, 2005; Golosov et al., 2006).

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 are as patient as 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 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 are less patient 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 a 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. Persson and Tabellini (1994) study a political model of capital taxation and show that necessary commitment under representative democracy corresponds closely to that provided by the actual institutions of most democracies. Bassetto (1996) explores how to sustain debt in an economy of renters and voters.

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 et al. (2008, 2010). Acemoglu et al. (2008) 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 no restrictions on tax policies. Acemoglu et al. (2010) use this framework for the analysis of the political economy of taxation and dynamic Mirrlees economies – the restrictions on taxes in that paper are endogenous and result from incentive compatibility constraints due to incomplete information. In our paper, we focus on the canonical Ramsey setup, where government is limited to linear (distortionary) taxes.

Most closely related to our paper is the recent work by Yared (2010), 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. In a political economy setup similar to ours, Caballero and Yared (2010) also study the dynamics of taxes, though they focus on a stochastic environment with aggregate shocks and ignore the role of capital taxation.

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.³ Albanesi and Armenter (2007a,b) provide a unified framework for the study of intertemporal distortions, though their framework does not incorporate explicit political economy considerations or allow the planner (politicians) and the agents to have different discount factors. Aguiar and Amador (2009) provide a tractable model for the effects of dynamic political economy on policy and capital accumulation. Several papers study Markov perfect equilibria in models of dynamic fiscal policy with time inconsistency or with political economy elements. Hassler et al. (2008), for example, show the possibility of positive long-run taxation and cycles in an environment with age-dependent capital depreciation rates. Aguiar et al. (2007, 2009) characterize optimal taxes and debt policy in a small open economy. Hassler et al. (2005), Song et al. (2009) and Battaglini and Coate (2008) study dynamic taxation in the presence of different political economy elements. Armenter (2007) shows that in a two-class, stochastic economy similar to that in Judd (1985), the standard Ramsey policy sequence can be sustained if policy revisions require unanimity to be approved. Farhi and Werning (2008) and Sleet and Yeltekin (2006, 2008) study dynamic fiscal policy in an environment with private information and lack of commitment or political economy constraints, and show that constrained optimal policies in these environments can be characterized as a solution to an

² Our focus on the best SPE is motivated by our attempt to 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.

³ There is also a large quantitative literature on time-inconsistent tax policies with benevolent politicians (social planners). For example, Klein et al. (2008) focus on time consistent Markovian equilibria, while Phelan and Stacchetti (2001) study more general sustainable equilibria in such environments.

optimal planning problem with a discount factor greater than the true discount factor.

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 (1999). We then introduce the political economy constraints.

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

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

where $c \geq 0$ denotes consumption, $l \geq 0$ is labor supply, and $\beta \in (0, 1)$ is the discount factor of the citizens. 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 and 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 $\lim_{c \rightarrow \infty} u'(c) = 0$.

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 , a nonnegative amount of public goods g_t needs to be financed, otherwise, the utility of the households is arbitrarily low (or equal to $-\infty$).⁴ 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 the wage rate and the interest rate (which is also the rental rate of capital) at time t , w_t and r_t , satisfy

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

The only tax instruments available to the government are linear taxes on capital, $\tau_{k,t} \leq 1$, and labor income, $\tau_{l,t} \leq 1$. 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 J 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

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

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 supplies 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 (arbitrarily low utility for the households). Let $b_t \in \mathbb{R}$ be the debt level of the government at time t (we restrict $b_0 = 0$), $q_{t+1} \geq 0$ 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 :

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

The left-hand side of Eq. (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 Eq. (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).

At any point in 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.

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 and labor, $\tau_{k,t}$ and $\tau_{l,t}$, respectively (with $0 \leq \tau_{k,t}, \tau_{l,t} \leq 1$), 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

⁴ More rigorously, we could define the utility function of each consumer as $u(c_t, \gamma_t)$, where $\gamma_t = 1$ denotes that the public good is supplied at time t . We do not do so to simplify the notation.

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

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 $\{\tau_{k,t}, \tau_{l,t}, q_t, x_t, \nu_t, \gamma_t, q_{t+1}\}$,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

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

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 . As in Chari and Kehoe (1993) and in Yared (2010), we impose that the households choose debt b_t in a bounded interval that can be set arbitrarily large. This ensures that no Ponzi condition is satisfied both on and off the equilibrium path.

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.

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).⁷ Note that consumers are anonymous and non-strategic in their private market behavior, though the representative citizen is strategic in his decision of whether to replace the current politician. Because households are anonymous, the public history h^t does not contain information on individual actions and public decisions are not conditioned on these. The politician in power is strategic in his choice of policies. 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 .

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 Eq. (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 present. 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 greater inefficiency than the best SPE. In particular, we refer to a SPE by the along-the-equilibrium path actions, that is, as $\{\tau_{k,t}, \tau_{l,t}, x_t, \nu_t, \gamma_t, \rho_t, c_t, l_t, b_t, q_{t+1}, k_{t+1}\}_{t=0}^{\infty}$.

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 $\{\tau_{k,t}, \tau_{l,t}, x_t, \nu_t, \gamma_t, q_{t+1}\}_{t=0}^{\infty}$, 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 Eq. (1) subject to Eq. (5) given $\{\tau_{k,t}, \tau_{l,t}, x_t, \nu_t, \gamma_t, q_{t+1}\}_{t=0}^{\infty}$ 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 Eq. (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 .
- iv. (*feasibility*) the feasibility constraint

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

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

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

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 $\nu_t = 1$, the value of holding capital and bonds must be the same, thus

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

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

Proof. This result follows from the definition of the competitive equilibrium, Definition 1, the conditions on factor prices (Eq. (2)), the first-order conditions on capital and labor (Eq. (7)), and the no-arbitrage condition (Eq. (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, $\{g_t\}_{t=0}^{\infty} = 0$, since this is already specified by the sequence $\{\tau_{k,t}, \tau_{l,t}, x_t, \nu_t, \gamma_t, \rho_t, c_t, l_t, b_t, q_{t+1}, k_{t+1}\}_{t=0}^{\infty}$), 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 (Eq. (7)) and $\{q_{t+1}\}_{t=0}^{\infty}$ must satisfy the no-arbitrage condition (Eq. (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 Eqs. (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. \square

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.,

⁶ Throughout, we refer to the tuple $\{\tau_{k,t}, \tau_{l,t}, q_t, x_t, \nu_t, \gamma_t, q_{t+1}\}$ 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.

⁷ For a standard treatment of SPE in a game between a government and a continuum of citizens, see Chari and Kehoe (1990). A full definition of an SPE is more involved than what we state in the text, since it requires that we specify the equivalents of the government budget constraint (4) and the implementability constraint (9) below for arbitrary histories. Yared (2010) provides a full definition of an SPE in a related model which can also be directly applied here. We omit the details to economize on space.

Atkinson and Stiglitz, 1980, Chari and Kehoe, 1999, 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 $k_0 \geq 0$. 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

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

Proof. Substitute the necessary and sufficient first-order conditions for utility maximization given in Eq. (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. \square

For our further analysis it is useful to point out that not all sequences $\{x_t, g_t\}_{t=0}^\infty$ are consistent with the existence of a competitive equilibrium. A competitive equilibrium does not exist if the present value of expenditures $x_t + g_t$ exceeds the present value of the maximal tax revenues. We define set Γ as a set of all sequences $\{x_t, g_t\}_{t=0}^\infty$ for which a competitive equilibrium exists. We call a sequence $\{x_t, g_t\}_{t=0}^\infty$ feasible if it satisfies

$$\{x_t, g_t\}_{t=0}^\infty \in \Gamma. \quad (10)$$

We denote the interior of the set Γ by $\text{Int}(\Gamma)$.

Given Proposition 2, the traditional analysis of optimal fiscal policy would proceed 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, if we did not raise such resources and set $x_t = 0$ for all t , the politician in power would be better off by taxing capital and labor at a very high rate and consuming the proceeds even if this meant being ousted from power. 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 gives us clues about the form of these sustainability constraints should take. At any point in time, the politician in power can deviate, collect all production as tax revenue, and consume all the proceeds. More specifically, if government owns debt, $b_t > 0$, the politician defaults, sets $\iota_t = \gamma_t = q_{t+1} = 0$, and $\tau_{k,t} = \tau_{l,t} = 1$, so that his consumption x_t is equal from Eq. (4) to

$$x_t = r_t K_t + w_t L_t = F(K_t, L_t).$$

If $b_t < 0$, politician still chooses $\gamma_t = q_{t+1} = 0$, and $\tau_{k,t}$ and $\tau_{l,t}$ are set to collect the maximum revenues while satisfying constraint (4). In particular, they are set to any level that satisfies

$$\tau_{k,t} r_t K_t + \tau_{l,t} w_t L_t - b_t = F(K_t, L_t).$$

The worst subgame perfect punishment from the viewpoint of the politician in power involves the citizens replacing this politician. After replacement, we assume that the politician receives zero consump-

tion 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

$$\sum_{s=0}^\infty \delta^s v(x_{t+s}) \geq v(F(k_t, l_t)). \quad (11)$$

We next show that Eq. (11) is in fact the relevant sustainability constraint. In particular, the next proposition proves that if the best allocation subject to Eq. (11) involves the provision of the public good in all periods, then the best SPE involves no political replacement (i.e., the initial politician will remain in power forever) and no default, and can be characterized as a solution to a simple maximization problem with Eq. (11) as an additional sustainability constraint.

Proposition 3. Suppose that given the sequence $\{g_t\}_{t=0}^\infty$, any solution to the maximization of Eq. (1), subject to the feasibility constraints, (6) and (10), the implementability constraint (9), and the political sustainability constraint (11) involves provision of the public good, $\gamma_t = 1$. Then, the best SPE $\{\tau_{k,t}^*, \tau_{l,t}^*, b_t^*, \iota_t^*, \gamma_t^*, \rho_t^*, x_t^*, c_t^*, l_t^*, k_{t+1}^*, q_{t+1}^*\}$ involves no political replacement, the required public good provision in all periods, and no default at all times (that is, $\rho_t^* = 0$ and $\gamma_t^* = \iota_t^* = 1$ for all t) along the equilibrium path. This best SPE can be characterized as maximizing the utility of the citizens (Eq. (1)), subject to the feasibility constraints (6) and (10), the implementability constraint (9), and the political sustainability constraint (11).

Proof. First, we show that no default occurs along the equilibrium path in the best SPE. This follows, since if $\iota_t^* = 0$ and $b_t^* > 0$ at some t (if $b_t^* \leq 0$, $\iota_t = 0$ is not allowed), then there would exist 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 would then imply $\iota_t^* = 1$. This contradiction establishes that $\iota_t^* = 1$ for all t . Then the best SPE involves public good provision at all dates is also straightforward by the hypothesis of the proposition (that any solution to maximizing Eq. (1), subject to Eqs. (6), (9), (10), and (11) involves $\gamma_t^* = 1$).

Since $\gamma_t^* = \iota_t^* = 1$, the best SPE satisfies the conditions in Proposition 2, and thus Eqs. (6), (9) and (10). Also note that by the argument preceding the sustainability constraint (11), this equation is a necessary condition, since otherwise the politician can improve his utility by deviating.

We next prove that $\rho_t^* = 0$ for all t , i.e., no political replacement along the equilibrium path. Suppose that there exists a best SPE that implements the maximization of Eq. (1), subject to Eqs. (6), (9), (10), and (11). 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 $\rho_t^* = 0$ so the best SPE involves no political replacement along the equilibrium path.

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 Eq. (1), subject to Eqs. (6), (9), (10), and (11). This implies that at t , $\gamma_t^* = \iota_t^* = 1$ and the politician's sustainability constraint, Eq. (11), 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

$$V(k_t^*) \equiv v(x_t^*) + \delta V(k_{t+1}^*) \geq v(F(k_t^*, l_t^*)), \quad (12)$$

⁸ The 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.

where the first relation is just a definition, and the inequality is imposed by Eq. (11). 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 a sequence of $\{\hat{x}_{t+s}\}_{s=1}^{\infty}$ and the continuation utility is

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

so that the sustainability constraint (11) for this new politician is satisfied. Now consider the following variation: do not replace the initial politician at 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 Eq. (11) after 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 Eq. (12) combined with the fact $V^R(k_{t+1}^*) > V(k_{t+1}^*) = 0$. But this implies that with this variation, the sustainability constraint, (11), 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 $\{\tau_{k,t}^*, \tau_{l,t}^*, b_t^*, \gamma_t^*, \rho_t^*, x_t^*, c_t^*, k_{t+1}^*, q_{t+1}^*\}$ could not have been a solution to the problem of maximizing Eq. (1), subject to Eqs. (6), (9), (10), and (11), yielding a contradiction and establishing the claim that the best SPE must involve $\rho_t^*=0$ for all t .

To complete the proof, we only need to show that the maximization of Eq. (1), subject to Eqs. (6), (9), (10) and (11) 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 defaults on government debt (provided that $b_t > 0$), does not finance g_t , and always chooses taxes $\tau_{k,t}$ and $\tau_{l,t}$ to consume all the output in the economy, $F(K_t, L_t)$. 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 Eq. (1), subject to the feasibility constraints, (6) and (10), the implementability constraint (9), and the political sustainability constraint (11) involves $\gamma_t=1$ for all t , that $\{g_t\}_{t=0}^{\infty}$ converges to some $g^S > 0$, and the best SPE equilibrium $\{\tau_{k,t}^*, \tau_{l,t}^*, b_t^*, \gamma_t^*, \rho_t^*, x_t^*, c_t^*, k_{t+1}^*, q_{t+1}^*\}$ 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) . Suppose that $\{x_t^*, g_t\}_{t=0}^{\infty} \in \text{Int}(\Gamma)$ and $c^S > 0$. Then we have that:*

1. if the politicians are as patient as the citizens, i.e., if $\delta = \beta$, then the sustainability constraint (11) 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 (11) binds as $t \rightarrow \infty$, and $\lim_{t \rightarrow \infty} \tau_{k,t}^* > 0$.

Proof. Since $\{x_t^*, g_t\}_{t=0}^{\infty} \in \text{Int}(\Gamma)$, constraint (10) does not bind and the sequence $\{c_t^*, l_t^*, b_t^*, k_{t+1}^*\}_{t=0}^{\infty}$ is a solution to maximization of Eq. (1) subject to Eqs. (6), (9) and (11). 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 (11).

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

$$u'(c_t^*) + \eta(u'(c_t^*) + u''(c_t^*)c_t^*) = \lambda_t, \tag{13}$$

$$h'(l_t^*) + \eta(h'(l_t^*) + h''(l_t^*)l_t^*) + \beta^{-t} \psi_t v'(F(k_{t+1}^*, l_{t+1}^*)) = \lambda_t F_L(k_{t+1}^*, l_{t+1}^*), \tag{14}$$

$$\lambda_t = \lambda_{t+1} \beta F_K(k_{t+1}^*, l_{t+1}^*) - \beta^{-t} \psi_{t+1} v'(F(k_{t+1}^*, l_{t+1}^*)) F_K(k_{t+1}^*, l_{t+1}^*), \tag{15}$$

$$\lambda_t \beta^t = \sum_{s=0}^t \delta^{t-s} \psi_s v'(x_t^*). \tag{16}$$

Note that by definition, the multiplier on the implementability constraint, η , must be finite. From Eq. (13) it follows that there exists $\lim_{t \rightarrow \infty} \lambda_t = \lambda^S < \infty$, because $\lim_{t \rightarrow \infty} c_t^* = c^S > 0$ is assumed to exist.

(Part 1) First, suppose that the discount factors of the politician and the citizens are equal, $\delta = \beta$. Then, Eq. (16) 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 . 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 (11) is violated for sufficiently large t , unless $F(k_t^*, l_t^*) \rightarrow 0$ (i.e., unless $F(k^S, l^S) = 0$). But the latter would imply that γ_t goes to 0 in finite time (since $g^S > 0$). By hypothesis, the maximization of Eq. (1) subject to Eqs. (6), (9) and (11) 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 $t \rightarrow \infty$, Eq. (11) becomes asymptotically slack.

Let us next take the limit as $t \rightarrow \infty$ in Eqs. (13), (14) and (15). Using the fact that $\beta^{-t} \psi_t \rightarrow 0$, these imply

$$u'(c^S) + \eta(u'(c^S) + u''(c^S)c^S) = \lambda^S, \tag{17}$$

$$h'(l^S) + \eta(h'(l^S) + h''(l^S)l^S) = \lambda^S F_L(k^S, l^S), \tag{18}$$

$$\lambda^S = \lambda^S \beta F_K(k^S, l^S). \tag{19}$$

Eqs. (17) and (18) imply that $\lambda^S > 0$. To see this, recall that $\lambda^S \geq 0$, because it is the multiplier on the resource constraint. Suppose that $\lambda^S = 0$. Then, since $h' > 0$ and $h'' > 0$, Eq. (18) implies that $\eta \in (-1, 0)$. However, since $u' > 0$ and $u'' < 0$, Eq. (17) cannot be satisfied with $\eta \in (-1, 0)$ and $\lambda^S = 0$. This yields a contradiction and establishes that $\lambda^S > 0$. In view of this, Eq. (19) implies that

$$\beta F_K(k^S, l^S) = \lim_{t \rightarrow \infty} \beta F_K(k_t^*, l_t^*) = 1. \tag{20}$$

Then, Eq. (7) combined with Eq. (20) implies that $\lim_{t \rightarrow \infty} \tau_{k,t}^* = 0$, completing the proof of Part 1 when $\delta = \beta$.

(Part 2). Now consider the case where $\delta < \beta$. By the hypothesis that a steady state exists, Eq. (13) implies that $\lambda_t \rightarrow \lambda^S$. First, to obtain a contradiction, suppose that $\lambda^S = 0$. From Eq. (16), 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 $t \rightarrow \infty$, Eqs. (17) and (18) 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 Eq. (16), 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 (13), (15) and (20) 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. \square

This proposition is the main result of our paper. The intuition for this result is that, when $\delta = \beta$, the political sustainability constraints are present, but the best SPE involves backloading of the payments to politicians.⁹ This backloading ensures that the sustainability constraint of the politician will ultimately 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.

We next provide an intuition for why the sustainability constraint of the politician is asymptotically slack. Suppose that $\delta = \beta$ and recall that the best equilibrium from the viewpoint of the citizens involves backloading the rewards to the politician and thus the utility given to the politician in power is (ultimately) increasing over time. Moreover, we have that $x_t^* \rightarrow x^S$. Suppose first that x_t^* converges to x^S in finite time, say at time $T < \infty$. This means that to prevent deviations at times $t < T$, the politician is being promised a rent stream equal to x^S at all times $t \geq T$. This also implies that if the sustainability constraint at times $t > T'$ for some $T' < \infty$ sufficiently large (in particular greater

than T) were removed, the same rent stream would be chosen to provide him with incentives at times $t \leq T'$. But this in turn implies that all of the sustainability constraint after T' are redundant and thus have zero Lagrange multipliers. By implication, there is no need to distort allocations to relax these sustainability constraints. The intuition for the case in which T is infinite is similar.

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 $\delta < \beta$. In this case, the sustainability constraint, (11), binds asymptotically. This implies that higher output must be associated with greater rents to politicians, since otherwise the politician would have an incentive to deviate. Therefore, there is an additional (opportunity) cost of increasing output for the citizens—the higher rents that need to be paid to the politician to prevent him from deviating given the higher output level. This reasoning in turn implies that reducing the capital stock away from the “first-best” level weakens the politician’s incentive to deviate and enables the citizens to reduce politician rents. Consequently, the best SPE is implemented by positive long-run capital taxes to keep the capital stock below its first-best level. In particular, if the economy had $\tau_k = 0$, Eq. (7) implies that each individual would choose the undistorted level of savings, leading either to the violation of the sustainability constraint or to higher rents for the 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.

Both parts of Proposition 4 are important. The first part suggests that the conclusions of the existing literature that the capital tax is zero may have a wider applicability than the commonly-used framework with a benevolent government. In particular, this result applies, as in our paper, to a class of environment in which the government is controlled by self-interested politicians without the ability to commit to future taxes. The second part might ultimately be the more important result, however. It introduces a new reason for positive equilibrium taxes on capital even in the long run when politicians are more impatient (short-sighted) than the citizens. This might be a better approximation to reality, particularly when there are exogenous reasons for which politicians lose power (even if they do not deviate from the prescribed sequence of actions). The second part of Proposition 4 thus 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 detailed calibration, but to give further intuition for the theoretical results derived in the previous section and to provide some simple insights about convergence to the steady state and the structure of taxes before such convergence takes place.

We choose standard functional forms. 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 $\beta = 0.95$.

⁹ See Acemoglu et al. (2008) for further discussion of backloading in political economy environments, and in particular, of the definition of “backloading” when there is the additional state variable given by the capital stock. See also Ray (2002) for a general treatment of backloading results in principal-agent models.

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 have 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 $g=0.05$ in each period. Fig. 1 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 Fig. 1 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 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.

We have also explored (in the Numerical Appendix, available upon request) a variety of other discount factors for the politicians ranging from $\delta=0.8$ to $\delta=0.95$. We briefly report the results here. The tax on capital τ^k is closely linked to the discount factor of the politician. When the politician is more impatient, he needs to be compensated with greater rents both in the short run and the long run, and this implies that the long-run tax rate on capital needs to be higher—mainly to reduce the long-run capital stock and relax the sustainability constraint of the politician. The tax on labor τ^l also changes in tandem with the tax rate on capital and is higher when the politician is more impatient—partly to finance, together with the tax on capital, the rents being paid to the politician.

4. Conclusion

The main result of the Ramsey paradigm of dynamic optimal taxation, first arrived by Chamley (1986) and Judd (1985), is that

¹⁰ To 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 .

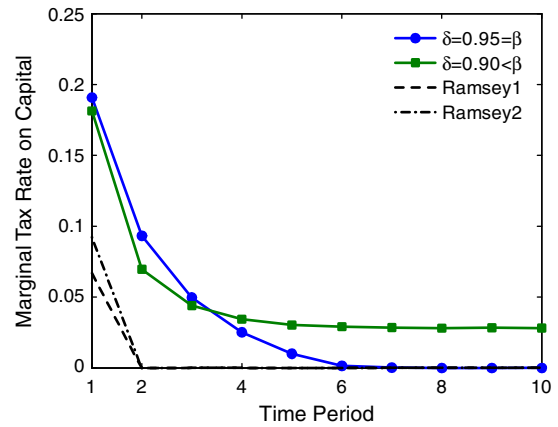


Fig. 1. The best SPE and Ramsey equilibria for different values of δ .

long-run capital taxes should be equal to zero. In practice, most societies have positive taxes on capital income. One perspective, adopted for example by Atkeson et al. (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 increases 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 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. But more importantly, it also emphasizes that 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 dynamic taxation in general, and capital taxation in particular, is another interesting area for future research.

Acknowledgments

We thank the editor and the anonymous referees for the useful comments. We are grateful to Andrew Atkeson, Tim Besley, V.V. Chari, Stephen Coate, and, especially, Pierre Yared for comments. We thank Georgy Egorov and Oleg Itskhoki for research assistance and the National Science Foundation for financial support. Golosov and Tsyvinski thank Einaudi Institute for Economics and Finance for hospitality.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:[10.1016/j.jpubeco.2010.10.014](https://doi.org/10.1016/j.jpubeco.2010.10.014).

References

- Abreu, Dilip, 1988. On the theory of repeated games with discounting. *Econometrica* 56 (2), 383–396.
- Acemoglu, Daron, Golosov, Mikhail, Tsyvinski, Aleh, 2008. Political economy of mechanisms. *Econometrica* 76 (3), 619–642.
- Acemoglu, Daron, Golosov, Mikhail, Tsyvinski, Aleh, 2010. Dynamic Mirrlees taxation under political economy constraint. *Review of Economic Studies* 77 (3), 841–881.
- Aguiar, Mark, Amador, Manuel, 2009. Growth in the shadow of expropriation, Stanford mimeo.
- Aguiar, Mark, Amador, Manuel, Gopinath, Gita, 2007. Efficient expropriation: sustainable fiscal policy in a small open economy. Unpublished.
- Aguiar, Mark, Amador, Manuel, Gopinath, Gita, 2009. Investment cycles and sovereign debt overhang. *Review of Economic Studies* 76 (1), 1–31.
- Albanesi, Stefania, Armenter, Roc, 2007a. Intertemporal distortions in the second best. NBER Working Paper 13629. CEPR Discussion Paper 6577.
- Albanesi, Stefania, Armenter, Roc, 2007a. Understanding Capital Taxation in Ramsey Models. Unpublished.
- Armenter, Roc, 2007. Time-consistent fiscal policy and heterogeneous agents. *Review of Economic Dynamics* 10, 31–54.
- Atkeson, Andrew, Chari, V.V., Kehoe, Patrick, 1999. Taxing capital income: a bad idea. *Federal Reserve Bank of Minneapolis Quarterly Review* 23, 3–17.
- Atkinson, Anthony, Stiglitz, Joseph, 1980. *Lectures in Public Economics*. McGraw-Hill, London.
- Barro, Robert, 1973. The control of politicians: an economic model. *Public Choice* 14 (1), 19–42.
- Bassetto, Marco, 1996. Optimal Fiscal Policy with Heterogenous Agents. unpublished.
- Battaglini, Marco, Coate, Stephen, 2008. A dynamic theory of public spending, taxation and debt. *American Economic Review* 98 (1), 201–236.
- Becker, Gary, Mulligan, Casey B., 2003. Deadweight costs and the size of government. *Journal of Law and Economics* 46 (2), 293–340.
- Benhabib, Jess, Rustichini, Aldo, 1997. Optimal taxes without commitment. *Journal of Economic Theory* 77 (2), 231–259.
- Besley, Timothy, Coate, Stephen, 1998. Sources of inefficiency in a representative democracy: a dynamic analysis. *The American Economic Review* 88 (1), 139–156.
- Besley, Timothy, Smart, Michael, 2007. Fiscal restraints and voter welfare. *Journal of Public Economics* 91 (3–4), 755–773.
- Brennan, Geoffrey, Buchanan, James, 1980. *The Power to Tax*. Cambridge University Press.
- Buchanan, James M., Tullock, Gordon, 1962. *The Calculus of Consent*. University of Michigan Press, Michigan.
- Caballero, Ricardo, Yared, Pierre, 2010. Future rent-seeking and current public savings. *Journal of International Economics* 82, 1124–1136.
- Chamley, Christophe, 1986. Optimal taxation of capital income in general equilibrium and infinite lives. *Econometrica* 54 (3), 607–622.
- Chari, V.V., Kehoe, Patrick, 1990. Sustainable plans. *Journal of Political Economy* 98 (4), 783–802.
- Chari, V.V., Kehoe, Patrick, 1993. Sustainable plans and mutual default. *Review of Economic Studies* 60 (1), 175–195.
- Chari, V.V., Kehoe, Patrick, 1999. Optimal fiscal and monetary policy. In: Taylor, John B., Woodford, Michael (Eds.), *Handbook of Macroeconomics*, vol. 1C, Elsevier Science, North-Holland.
- Chari, V.V., Christiano, Lawrence, Kehoe, Patrick, 1994. Optimal fiscal policy in a business cycle model. *Journal of Political Economy* 102 (4), 617–652.
- Erosa, Andres, Gervais, Martin, 2002. Optimal taxation in life-cycle economies. *Journal of Economic Theory* 105 (2), 338–369.
- Farhi, Emmanuel and Ivan Werning, 2008. *The Political Economy of Nonlinear Capital Taxation*, mimeo, Harvard and MIT.
- 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, Mikhail, Kocherlakota, Narayana, Tsyvinski, Aleh, 2003. Optimal indirect and capital taxation. *Review of Economic Studies* 70 (3), 569–587.
- Golosov, Mikhail, Tsyvinski, Aleh, Werning, Ivan, 2006. New dynamic public finance: a user's guide. NBER Macroeconomics Annual 2006. MIT Press, Cambridge, MA.
- Hassler, John, Krusell, Per, Storesletten, Kjetil, Zilibotti, Fabrizio, 2005. The dynamics of government. *Journal of Monetary Economics* 52 (7), 1331–1358.
- Hassler, John, Krusell, Per, Storesletten, Kjetil, Zilibotti, Fabrizio, 2008. On the timing of the optimal capital taxation. *Journal of Monetary Economics* 55 (4), 692–709.
- Jones, Larry E., Manuelli, Rodolfo E., Rossi, Peter E., 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, Krusell, Per, Ríos-Rull, José-Víctor, 2008. Time consistent public expenditures. *Review of Economic Studies* 75 (3), 789–808.
- Kocherlakota, Narayana, 2005. Zero expected wealth taxes: a Mirrlees approach to dynamic optimal taxation. *Econometrica* 73 (5), 1587–1621.
- Ljungqvist, Lars, Sargent, Thomas J., 2004. *Recursive Macroeconomic Theory*. MIT Press, Cambridge, MA.
- Persson, Torsten, Tabellini, Guido, 1994. Representative democracy and capital taxation. *Journal of Public Economics* 55 (1), 53–70.
- Persson, Torsten, Tabellini, Guido, 2000. *Political Economics: Explaining Economic Policy*. MIT Press, Cambridge, MA.
- Persson, Torsten, Tabellini, Guido, 2004. Constitutional rules and fiscal policy outcomes. *The American Economic Review* 94 (1), 25–45.
- Phelan, Christopher, Stacchetti, Ennio, 2001. Sequential equilibria in a Ramsey tax model. *Econometrica* 69 (6), 1491–1518.
- Ray, Debraj, 2002. Time structure of self-enforcing agreements. *Econometrica* 70 (2), 547–582.
- Reis, Catarina, 2007. Taxation without Commitment. Unpublished.
- Sleet, Christopher, Yeltekin, Sevin, 2006. Credibility and endogenous societal discounting. *Review of Economic Dynamics* 9, 410–437.
- Sleet, Christopher, Yeltekin, Sevin, 2008. Politically credible social insurance. *Journal of Monetary Economics* 55, 129–151.
- Song, Zheng, Storesletten, Kjetil, Zilibotti, Fabrizio, 2009. Rotten Parents and Disciplined Children: A Politico-Economic Theory of Public Expenditure and Debt. Working paper.
- Wilson, John D., 1989. On the optimal tax base for commodity taxation. *The American Economic Review* 79 (5), 1196–1206.
- Yared, Pierre, 2010. Politicians, taxes, and debt. *Review of Economic Studies* 77, 806–840.
- Zhu, Xiaodong, 1992. Optimal fiscal policy in a stochastic growth model. *Journal of Economic Theory* 58 (2), 250–289.