# Sequential Equilibria in a Ramsey Tax Model

Christopher Phelan
*Federal Reserve Bank of Minneapolis*

and

Ennio Stacchetti
*University of Michigan*

**First Draft**: November 18, 1996
**This Version**: September 28, 2000

# 1. Introduction

This paper presents a full characterization of the equilibrium value set of a Ramsey tax model where the government cannot commit to a sequence of tax rates. More generally, it develops a dynamic programming method for a class of policy games between the government and a continuum of households.

Starting with the seminal paper of Kydland and Prescott (1977), there is by now a large body of literature dealing with reputation and credibility of government policies (see, for example, Chari and Kehoe (1990, 1993a, 1993b), Persson, Persson and Svensson (1987), Rogoff (1989), and Stokey (1989, 1991)). One of the main thrusts of the paper by Kydland and Prescott (1977) was that the inability of governments to commit to policies made the application of optimal control techniques to questions of government policy *inappropriate.* Later papers (such as Chari and Kehoe (1990)) argue that a policy problem is better viewed as a dynamic game between the government and a continuum of households. The strategic dynamic programming approach of Abreu, Pearce and Stacchetti (1986, 1990) is a natural starting point to study these policy games. Atkeson (1991) has already demonstrated that strategic dynamic programming can incorporate the presence of a *public* state variable (in his case, the capital held by the single borrower). If capital levels are *privately* held by anonymous households, however, tractability becomes a major concern.

In a model with a continuum of households, a direct extension of Abreu, Pearce and Stacchetti's approach to the macro policy game requires that each household be assigned a continuation value after each history. These values provide incentives for conforming to equilibrium behavior, even once a household has the "wrong" level of personal assets because it did not conform in the past. This produces a problem that is prohibitively complex, either for the derivations of interesting properties or for numerical computations.

We propose instead a much simpler accounting system, wherein we only keep track of the "marginal value of capital" for a household that has never deviated from the equilibrium, and invoke the "convexity" of the household's dynamic consumption problem to ignore totally the incentives of households that have previously deviated. Kydland and Prescott (1980) and Marcet and Marimon (1994) have demonstrated that keeping track of the marginal value of capital, the Ramsey problem with commitment has a recursive structure. We show that a similar result attains in a strategic setting without commitment. The crucial observation is that since each agent (other than the government) is anonymous and cannot affect the path of prices, the household's problem, unlike that of the government, *can* be viewed as a standard optimal control problem. By selectively incorporating Euler conditions into the strategic dynamic programming framework, we wed two technologies that are usually considered competing alternatives.

Though we develop our method for a specific model, the technique should be useful for a class of policy games. We consider a one-sector growth model with capital, labor, and a public good that must be financed by distortionary taxes on capital and labor. The government is unable to commit to future tax rates. Any tax policy by the government produces a dynamic economy with a corresponding competitive equilibrium.

In a similar model, Chari and Kehoe (1990) propose a method for checking whether

an arbitrary tax policy is the outcome of a symmetric sequential equilibrium:[1] in every period the government should not be able to improve the total welfare of a representative household by changing the current period's tax rates, if this is then followed by the worst sequential equilibrium of the ensuing subgame. However, this approach assumes that one can easily construct a worst equilibrium. For many interesting models (in particular for ours), a worst equilibrium is not readily available. Their method focuses mainly on the best sustainable tax policy and provides little guidance on how to generate candidate tax policies. In contrast, our method deals directly with the entire set of equilibrium values (and so it does not require that the worst equilibrium be known in advance) and delivers a recursive algorithm for computing this set.

The characterization of the entire equilibrium value set also facilitates the examination of other questions. We study the steady state of the best equilibrium without commitment. A celebrated result of Chamley (1986) and Judd (1985) states that with full commitment, the optimal capital tax rate converges to 0 in the steady state. In our anonymous game, this may not be sustainable, and thus the steady state of the best equilibrium without commitment may have positive taxes on capital and may lead to a steady state capital $k^s$, different from that which would attain with full commitment. In this case, we show that the steady state corresponds to an extreme point of the set of equilibrium values associated with an initial capital $k^s$. In particular, it has the *lowest* equilibrium value (and either the lowest or the highest marginal value of capital) that can be attained from $k^s$. Moreover, the steady state path is optimal among all steady state paths that are restricted to maintaining the same steady state capital $k^s$. We present a computed example where a steady state with zero capital taxes is not sustainable in equilibrium and where the best equilibrium converges to a steady state with positive taxes on capital.

Three papers related to ours deserve special mention. In an independent work, Chang (1998) considers a monetary model and derives methods similar to ours. While his policy game is different, the two games share the common trait that households make anonymous decisions (and thus, unilateral deviations are not observable). In Chang (1988) the state variable is money, while here it is capital. Like us, Chang (1988) shows that equilibria can be characterized in terms of their value to the government and their marginal value of the private state variable (money). The methods differ in two respects. First, since the quantity of money, unlike capital, can be renormalized to unity in every period, his is a repeated game. Second, we exploit the ability to use public randomization in both the characterization and computation of equilibria.

The works of Benhabib and Rustichini (1997) and Marcet and Marimon (1994) propose an alternative method for studying policy without commitment. In particular, they argue that the optimal control approach to policy (as opposed to the game theoretic approach we employ) is indeed appropriate if additional constraints are imposed to ensure that the government does not deviate. These methods concentrate on the best policy without commitment and are particularly efficient if the worst punishment for a deviating government is known in advance.

We introduce the model and the assumptions in Section 2. Section 3 characterizes

---

[1]   Chari and Kehoe (1990) call such an equilibrium a *sustainable equilibrium.*

the competitive equilibria of the dynamic economy where the government tax policy is fixed (but not necessarily stationary). In Section 4 we define a sequential equilibrium for our anonymous game. In Section 5 we extend the dynamic programming tools developed by Abreu, Pearce and Stacchetti (1990) for repeated games. We thus are able to characterize the equilibrium value correspondence. In Section 6 we study the steady state properties of a best equilibrium of our no-commitment game. Section 7 presents an example, and Section 8 concludes.

## 2. The Game

We consider a dynamic game with a benevolent government and a continuum of identical households. Each household is endowed with 1 unit of labor-leisure every period. There is a single consumption good that is produced with a constant returns to scale technology. If capital and labor per capita are $K$ and $L$, then output per capita is $f(K, L)$. At the beginning of every period $t$, the outcome $x_t$ of a uniform $[0, 1]$ random variable $X_t$ is publicly observed. The random variables $\{X_t\}$ are serially uncorrelated and independent of any choices made by the government or the households. Next, the government chooses the capital tax rate $\tau_{k,t}$ and labor tax rate $\tau_{\ell,t}$ in the interval $[\underline{\tau}, \overline{\tau}]$. Finally, the households simultaneously choose their labor input and the fractions of their current income (from wages and capital rents) to consume and save (as capital) for production in the next period. Capital fully depreciates each period.[2] The totality of tax revenue is used by the government to finance a public good period by period, using a linear technology that transforms each unit of the consumption good into a unit of the public good.

If a household consumes the stream $\{(\ell_t, c_t, G_t)\}_{t=0}^{\infty}$ of labor, consumption good, and public good per capita, then its total discounted utility is

$$\sum_{t=0}^{\infty} \beta^t \left[ u(\ell_t, c_t) + g(G_t) \right],$$

where $\beta \in (0, 1)$ is the discount factor. All households have the same preferences and are strategically anonymous. That is, their individual actions cannot be observed, and the government can only react to the history of "sunspots" $x_t$, to its own past actions and to past distributions of consumption, capital and labor, all of which are publicly observed. The government's objective is to maximize the expected total discounted utility of a representative household.

Let $\Gamma(k_0)$ denote the dynamic game between the government and the households when all the households, except perhaps for a subset of measure 0, are endowed with initial capital stock $k_0$. As we explain later, since we restrict attention to symmetric strategy equilibria (where all households make the same choices along the equilibrium path), we will not need to explicitly consider situations where at the beginning of a period there is

---

[2] We make this assumption to keep the analysis simple; our results can be extended to the case where the depreciation rate is less than 1 and investments are irreversible. The example in Section 7 assumes that the depreciation rate is 0.05.

a nontrivial distribution of capital. Let $\Gamma(k_0, x_0)$ denote the subgame of $\Gamma(k_0)$ after the realization of $x_0$ but before the government sets period 0 taxes.

ASSUMPTIONS: The following assumptions will be in force throughout the paper:

(A1) $0 \leq \underline{\tau} < \overline{\tau} < 1$.

(A2) $g : \mathbf{R}_+ \to \mathbf{R}$ and $u : [0,1) \times \mathbf{R}_+ \to \mathbf{R}$ are continuously differentiable, $g$ is concave and increasing, and $u(\ell, c)$ is concave (jointly in $\ell$ and $c$), increasing in $c$, and decreasing in $\ell$. Moreover, $g(0) = 0$, and there exists $U_c \in \mathbf{R}$ such that for all $\ell \in [0,1)$ and $c \in \mathbf{R}_+$,

$$u_c(\ell, 0) \leq U_c, \quad \inf_{\ell' \in [0,1)} u_c(\ell', c) > 0, \quad \text{and} \quad u_\ell(0,0) = 0.$$

(A3) $f : \mathbf{R}_+^2 \to \mathbf{R}_+$ is concave and homogeneous of degree 1 and is continuously differentiable in the interior of $\mathbf{R}_+^2$. For all $(k, \ell) \in \operatorname{int}(\mathbf{R}_+^2)$, $f(0, \ell) = f(k, 0) = 0$, $f_k(k, \ell) > 0$ and $f_\ell(k, \ell) > 0$. Moreover, for all $\ell \in [0, 1]$,

$$\lim_{k \to 0} f_k(k, \ell) = \infty \quad \text{and} \quad \lim_{k \to \infty} f_k(k, \ell) = 0.$$

Most of these assumptions are standard for the one-sector growth model. Assumptions (A3), for instance, imply that competitive prices are equal to marginal products, that firms make zero profits, and that there exists a maximal sustainable capital stock $\overline{k}$ such that $\overline{k} = f(\overline{k}, 1)$ and $k > f(k, 1)$ for all $k > \overline{k}$. Maximal tax rates are bounded away from 100% in order to ensure that capital remains bounded away from zero.[3] The condition $u_c(\ell, 0) \leq U_c$, however, is not standard. We make it to guarantee that a household's transversality condition is satisfied (see Section 3).[4]

An action for the government at date $t$ is simply a pair of tax rates $\tau_t = (\tau_{k,t}, \tau_{\ell,t}) \in [\underline{\tau}, \overline{\tau}]^2$. An action $a_t = (\ell_t, \theta_t) \in [0,1]^2$ for a household at date $t$ includes a labor choice $\ell_t$ and a savings choice (as a fraction of current income) $\theta_t$. Let $K_t$ and $L_t$ be the average levels of capital and labor, and denote $p_{k,t} = f_k(K_t, L_t)$ and $p_{\ell,t} = f_\ell(K_t, L_t)$. The per capita government tax revenue is

$$G_t = \tau_{k,t} p_{k,t} K_t + \tau_{\ell,t} p_{\ell,t} L_t,$$

---

[3] We do not allow $\underline{\tau} < 0$. In our model the government cannot borrow or lend, and thus the government's revenue must be nonnegative in every period. Hence, if $\underline{\tau} < 0$, the government could, for example, subsidize labor with revenues from the capital tax. However, this introduces an additional difficulty, which we have preferred to avoid. Since tax rates are announced *before* the households make their decisions, it is possible that when the government is prepared to subsidize labor (or capital), it doesn't collect enough tax revenue to keep up its promise. Thus, to allow for $\underline{\tau} < 0$, we would have to choose some form of rationing for those contingencies. This complicates the analysis considerably.

[4] This rules out some commonly used utility functions. However, the lemmas of Section 3, including the appropriate transversality condition, also hold under different assumptions. For example, we can show that their conclusions are valid for the following pair of utility and production functions:

$$u(\ell, c) = v(\ell) + \frac{c^{1-\psi}}{1-\psi} \quad (0 < \psi < 1), \quad f(k, \ell) = A[\alpha k^\rho + (1-\alpha)\ell^\rho]^{\frac{1}{\rho}} \quad (A > 0, 0 < \alpha < 1, \rho \leq 0),$$

where $v : [0,1) \to \mathbf{R}$ is a decreasing, concave function, with $\lim_{\ell \to 1} v(\ell) = -\infty$. Note that when $\rho = 0$, $f(k, \ell) = A k^\alpha \ell^{1-\alpha}$.

and the income for a household with capital $k_t$ that chooses labor $\ell_t$ is

$$y_t = (1 - \tau_{k,t})p_{k,t}k_t + (1 - \tau_{\ell,t})p_{\ell,t}\ell_t.$$

If the household saves a fraction $\theta_t$, its consumption and capital investment are $c_t = (1 - \theta_t)y_t$ and $k_{t+1} = \theta_t y_t$, respectively. Though the agents' payoffs in $\Gamma(k_0)$ are affected by the initial capital stock $k_0$, the actions available to the government and the households are not. That is, the action spaces for the government and the agents in every period are independent of the current capital stocks.

With full generality, a public history at date $t$ should include the history of random outcomes $x^t = \{x_0, \ldots, x_t\}$, the history of taxes $\tau^t$, and the observed *distributions* of household actions $a^t$. We do not, however, need to specify play after all such histories. In this paper, we study exclusively *symmetric* strategy profiles, where all households choose the same actions along the path of play. Therefore, unless a positive measure of households deviate, all observed distributions of household actions are degenerate along the path of play. Since our analysis concerns sequential equilibria only, behavior after simultaneous deviations is irrelevant for checking the incentives of the players, and thus does not need to be specified. Hence, we define a public history at the end of period $t$ as a sequence $h^t = \{h_0, \ldots, h_t\}$ where $h_s = (x_s, \tau_s)$. Let $H = [0, 1] \times [\underline{\tau}, \overline{\tau}]^2$. Then, for any $t \geq 0$, $H^{t+1}$ is the set of all possible public histories $h^t$. (Define $H^0 = \{\emptyset\}$.)

A household's action in period $t$ should be a function of the public history $h^t$ as well as its private history of play $a^{t-1}$, which determines its date $t$ level of capital. However, the purpose of specifying off-path play is to ensure that a household has the appropriate incentives to follow the equilibrium. We argue below that, unlike in more general games, the convexity of the household's problem allows us to check for the optimality of on-the-path household play without specifying the household's off-the-path behavior. Therefore, explicitly specifying such off-path behavior is not necessary, and thus we make household strategies a function of the public history only.

A strategy for the government is a function $\sigma_G$ specifying a pair of tax rates $\tau_t = \sigma_G(t)(h^{t-1}, x_t)$ for each period $t$, history $h^{t-1}$, and random outcome $x_t$. A symmetric strategy for the households is a function $\sigma_C$ specifying an action $a_t = \sigma_C(t)(h^t)$ as a function of the public history $h^t = (h^{t-1}, x_t, \tau_t)$ after the government's decision. Both $\sigma_G(t)$ and $\sigma_C(t)$ are assumed to be measurable functions. A symmetric strategy profile for $\Gamma(k_0)$ is a pair of strategies $\sigma = (\sigma_C, \sigma_G)$. Let $\Sigma = \Sigma_C \times \Sigma_G$ denote the set of all symmetric strategy profiles for $\Gamma(k_0)$.

## 3. COMPETITIVE EQUILIBRIA

In this section we characterize the competitive equilibria of the dynamic economy in which a tax policy is arbitrarily specified. Although in our game the government does not have this commitment power, along the equilibrium path, the households act as if the government had established such a policy. Thus, an arbitrary stochastic policy $\tau = \{\tau_t(x^t)\}_{t=0}^{\infty}$ produces an auxiliary problem that is useful for characterizing the sequential equilibria of $\Gamma(k_0)$. Denote this economy $\Gamma(k_0 \mid \tau)$.

Let $q = \{(\ell_t, c_t, k_{t+1})(x^t)\}_{t=0}^{\infty}$ denote a competitive equilibrium of $\Gamma(k_0 \mid \tau)$ with corresponding (after tax) prices $p_{k,t}(x^t) = (1 - \tau_{k,t}(x^t))f_k(k_t(x^{t-1}), \ell_t(x^t))$ and $p_{\ell,t}(x^t) = (1 - \tau_{\ell,t}(x^t))f_\ell(k_t(x^{t-1}), \ell_t(x^t))$. Let

$$m_{t+1}(x^t) = E_{x_{t+1}} \Big[ p_{k,t+1}(x^{t+1})u_c(\ell_{t+1}(x^{t+1}), c_{t+1}(x^{t+1})) \mid x^t \Big]. \tag{1}$$

This quantity represents, in period $t$, the expected derivative of the houshold's lifetime discounted utility from period $t + 1$ on with respect to $k_{t+1}$.

Abreu, Pearce, and Stacchetti (1986, 1990) (hereafter APS) show that a sequential equilibrium for an infinitely repeated game can be reconstructed by piecing together the equilibria of a sequence of static games with appropriately augmented payoffs. A similar result holds for $\Gamma(k_0 \mid \tau)$. Consider the one-period economy where households have initial capital $k_t$, the government tax rates are $\tau_t$, and households' augmented utility function over consumption, labor, and end-of-period capital is $u(\ell_t, c_t) + \beta m_{t+1}k_{t+1}$, where $m_{t+1}$ is an exogenous parameter. Let $CE(k_t, \tau_t, m_{t+1})$ denote the set of competitive equilibrium allocations $(\ell_t, c_t, k_{t+1})$ of this static economy. In the rest of the paper, we exploit the fact that a necessary and sufficient condition for $q$ to be a competitive equilibrium of $\Gamma(k_0 \mid \tau)$ is that for all $t$ and $x^t$, $(\ell_t, c_t, k_{t+1})(x^t) \in CE(k_t(x^{t-1}), \tau_t(x^t), m_{t+1}(x^t))$, where $m_{t+1}(x^t)$ is defined by Equation (1).

The next three lemmas establish the sufficiency of this condition. We first find a uniform upper bound on $\{m_t(x^{t-1})\}$ for all policies $\tau$ and sequences $\{x_t\}$. This upper bound is used to show that a certain transversality condition is satisfied. To construct such a bound, we first need to find a uniform lower bound $\underline{k} > 0$ such that the resulting capital stock trajectory of the competitive equilibrium of $\Gamma(k_0 \mid \tau)$ is bounded below by $\underline{k}$ for all $k_0 \geq \underline{k}$ and all feasible tax policies $\tau$.

LEMMA 1: *There exists $\underline{k} > 0$ such that for all $k_0 \in [\underline{k}, \overline{k}]$ and tax policies $\tau$, all competitive equilibria of $\Gamma(k_0 \mid \tau)$ have $k_t(x^{t-1}) \geq \underline{k}$ for all $t \geq 0$ and $x^{t-1} \in [0,1]^t$.*

The proof of Lemma 1 is deferred to the Appendix. In what follows, we will only consider initial endowments above the minimum capital associated with competitive equilibria and below the maximal sustainable capital:

(A4) $k_0 \in [\underline{k}, \overline{k}]$.

Assumption (A4) ensures that if $k_0 \in [\underline{k}, \overline{k}]$, then for all feasible tax policies $\tau$ and all realizations $x^{t-1}$, $k_t(x^{t-1}) \in [\underline{k}, \overline{k}]$.

LEMMA 2: *There exists $\overline{m} < \infty$ such that for all $k_0 \in [\underline{k}, \overline{k}]$ and tax policies $\tau$, all competitive equilibria of $\Gamma(k_0 \mid \tau)$ satisfy $m_t(x^{t-1}) \leq \overline{m}$ for all $t \geq 0$ and $x^{t-1} \in [0,1]^t$.*

PROOF: Since $f(k, \ell) = f_k(k, \ell)k + f_\ell(k, \ell)\ell$, for all $t \geq 0$ and $x^t \in [0,1]^{t+1}$,

$$f_k(k_t, \ell_t) \leq \frac{f(k_t, \ell_t)}{k_t} \leq \frac{f(\overline{k}, 1)}{\underline{k}} = \overline{k}/\underline{k}.$$

By assumption (A2), $u_c(\ell_t, c_t) \leq U_c$. Hence, we can choose

$$\overline{m} := (1 - \underline{\tau})[\overline{k}/\underline{k}]U_c. \qquad \blacksquare$$

6

LEMMA 3: *Given a tax policy $\tau$ and an initial capital stock $k_0 \in [\underline{k}, \overline{k}]$, suppose that the sequence $q^* = \{(\ell_t^*, c_t^*, k_{t+1}^*)(x^t)\}_{t=0}^{\infty}$ is such that for all $t$ and $x^t$, $(\ell_t^*, c_t^*, k_{t+1}^*)(x^t) \in CE(k_t^*(x^{t-1}), \tau_t(x^t), m_{t+1}^*(x^t))$, where $m_{t+1}^*(x^t)$ is defined by Equation (1). Then $q^*$ (with prices $p_{k,t}^* = (1 - \tau_{k,t})f_k(k_t^*, \ell_t^*)$ and $p_{\ell,t}^* = (1 - \tau_{\ell,t})f_\ell(k_t^*, \ell_t^*))$ is a competitive equilibrium of $\Gamma(k_0 \mid \tau)$.*

PROOF: All we need to show is that $q^*$ is an optimal solution of the households' optimization problem (taking prices as given). Necessary conditions for $(\ell_t^*, c_t^*, k_{t+1}^*) \in CE(k_t^*, \tau_t, m_{t+1}^*)$ are

$$(u_\ell(\ell_t^*, c_t^*) + u_c(\ell_t^*, c_t^*)p_{\ell,t}^*)(\ell_t - \ell_t^*) \leq 0 \quad \text{for all } \ell_t \in [0, 1]$$
$$(-u_c(\ell_t^*, c_t^*) + \beta m_{t+1}^*)(k_{t+1} - k_{t+1}^*) \leq 0 \quad \text{for all } k_{t+1} \in [0, p_{k,t}^* k_t^* + p_{\ell,t}^* \ell_t^*].$$

These imply for all $t$ and $x^{t-1}$

$$E_{x_t}[u_\ell(\ell_t^*, c_t^*)(\ell_t - \ell_t^*) + u_c(\ell_t^*, c_t^*)[p_{\ell,t}^*(\ell_t - \ell_t^*) - (k_{t+1} - k_{t+1}^*)] \mid x^{t-1}]$$
$$\leq -E_{x_t}[\beta m_{t+1}^*(k_{t+1} - k_{t+1}^*) \mid x^{t-1}],$$

and if we add $E_{x_t}[u_c(\ell_t^*, c_t^*)p_{k,t}^*(k_t - k_t^*) \mid x^{t-1}] = E_{x_t}[m_t^*(k_t - k_t^*) \mid x^{t-1}]$ to both sides of this inequality, we obtain

$$E_{x_t}[u_\ell(\ell_t^*, c_t^*)(\ell_t - \ell_t^*) + u_c(\ell_t^*, c_t^*)(c_t - c_t^*) \mid x^{t-1}]$$
$$\leq E_{x_t}[m_t^*(k_t - k_t^*) - \beta m_{t+1}^*(k_{t+1} - k_{t+1}^*) \mid x^{t-1}].$$

Also, the concavity of $u$ implies that

$$u(\ell_t, c_t) \leq u(\ell_t^*, c_t^*) + u_\ell(\ell_t^*, c_t^*)(\ell_t - \ell_t^*) + u_c(\ell_t^*, c_t^*)(c_t - c_t^*).$$

Therefore

$$D := \lim_{T \to \infty} E\left[\sum_{t=0}^{T} \beta^t[u(\ell_t, c_t) - u(\ell_t^*, c_t^*)]\right]$$
$$\leq \lim_{T \to \infty} E\left[\sum_{t=0}^{T} \beta^t[m_t^*(k_t - k_t^*) - \beta m_{t+1}^*(k_{t+1} - k_{t+1}^*)]\right]$$
$$= \lim_{T \to \infty} E\left[\beta^{T+1} m_{T+1}^*(k_{T+1}^* - k_{T+1})\right] \leq \lim_{T \to \infty} E\left[\beta^{T+1} m_{T+1}^* k_{T+1}^*\right],$$

because $k_0^* = k_0$, and $m_{T+1}^*$ and $k_{T+1}$ are nonnegative. From Lemma 2 we have that $m_t^* \leq \overline{m}$ for all $t$ and $x^t$. Since $k_0^* \leq \overline{k}$, $k_t^* \leq \overline{k}$ for all $t$ and $x^t$. Therefore, $m_t^* k_t^* \in [0, \overline{m}\overline{k}]$ for each $t$ and $x^t$, and

$$\lim_{T \to \infty} E\left[\beta^T m_T^* k_T^*\right] = 0.$$

The latter is the appropriate transversality condition. Hence, $D \leq 0$, and the plan $q^*$ is optimal. ∎

## 4. Sequential Equilibria

We now provide a formal definition of a sequential equilibrium for our policy game. This is a direct adaptation of the same concept for standard dynamic games, and as stated in the Introduction, it is the same concept that Chari and Kehoe (1990) call a *sustainable equilibrium*.

A symmetric strategy profile $\sigma$ for $\Gamma(k_0)$ generates a unique random *outcome path* $\{(\tau_t, a_t)(x^t)\}_{t=0}^{\infty}$ and corresponding random allocation process $q = \{(\ell_t, c_t, k_{t+1})(x^t)\}_{t=0}^{\infty}$ as follows. Let $h^{-1} = \emptyset$, and for any sequence of random outcomes $\{x_s\}_{s=0}^{\infty}$, inductively define for all $t \geq 0$

$$\tau_t(x^t) = \sigma_G(t)(h^{t-1}(x^{t-1}), x_t), \qquad a_t(x^t) = \sigma_C(t)(h^{t-1}(x^{t-1}), x_t, \tau_t(x^t)),$$

$$y_t(x^t) = (1 - \tau_{k,t}(x^t))f_k(k_t(x^{t-1}), \ell_t(x^t))k_t(x^{t-1})$$
$$+ (1 - \tau_{\ell,t}(x^t))f_\ell(k_t(x^{t-1}), \ell_t(x^t))\ell_t(x^t),$$

$$c_t(x^t) = (1 - \theta_t(x^t))y_t(x^t), \qquad k_{t+1}(x^t) = \theta_t(x^t)y_t(x^t),$$

$$G_t(x^t) = f(k_t(x^{t-1}), \ell_t(x^t)) - y_t(x^t), \qquad h^t(x^t) = (h^{t-1}(x^{t-1}), x_t, \tau_t(x^t)).$$

A strategy profile $\sigma \in \Sigma$ induces after any history $h^{t-1} \in H^t$ a (symmetric) strategy profile $\sigma|_{h^{t-1}} \in \Sigma$. For all $s \geq 0$, $\hat{h}^{s-1} \in H^s$, $\hat{x}_s \in [0,1]$, and $\hat{\tau}_s \in [\underline{\tau}, \overline{\tau}]^2$,

$$\sigma_G|_{h^{t-1}}(s)(\hat{h}^{s-1}, \hat{x}_s) := \sigma_G(t+s)(h^{t-1}, \hat{h}^{s-1}, \hat{x}_s)$$

$$\sigma_C|_{h^{t-1}}(s)(\hat{h}^{s-1}, \hat{x}_s, \hat{\tau}_s) := \sigma_C(t+s)(h^{t-1}, \hat{h}^{s-1}, \hat{x}_s, \hat{\tau}_s).$$

That is, $\sigma|_{h^{t-1}}$ generates in period $s$, after history $\hat{h}^{s-1}$ and outcome $\hat{x}_s$, the same actions as the strategy $\sigma$ generates in period $t + s$, after the history $h^{t-1}$ followed by the history $\hat{h}^{s-1}$ and $\hat{x}_s$. We can similarly define $\sigma|_{(h^{t-1}, x_t)}$ as the strategy profile generated by $\sigma$ after history $h^{t-1}$ and realization $x_t$. In particular, for any $x_0 \in [0,1]$, $\sigma|_{x_0}$ is a symmetric strategy profile for $\Gamma(k_0, x_0)$. Sometimes we will also view $\sigma|_{x_0}$ as the symmetric strategy profile for $\Gamma(k_0)$ that prescribes $\sigma|_{x_0}$ after all realizations of $X_0$. That is, we can view $\sigma|_{x_0}$ as a strategy profile for $\Gamma(k_0)$ that prescribes deterministic actions for the first period.

The *value* of $\sigma \in \Sigma$ is the corresponding government's normalized[5] total discounted payoff:

$$\Phi_G(k_0, \sigma) := (1 - \beta)E\Big[\sum_{t=0}^{\infty} \beta^t \left[u(\ell_t, c_t) + g(G_t)\right]\Big].$$

As explained above, for any $x_0 \in [0,1]$, we can view $\sigma|_{x_0}$ as a strategy profile for $\Gamma(k_0)$. Then we can use the same formula above to compute $\Phi_G(k_0, \sigma|_{x_0})$. In addition,

$$\Phi_G(k_0, \sigma) = E_{x_0}\big[\Phi_G(k_0, \sigma|_{x_0})\big].$$

---

[5]   The normalization translates the discounted sum of payoffs into a *weighted average*, with weights $(1 - \beta)\beta^t$ that add up to 1. This makes the value of $\sigma$ comparable with the payoffs received in any period and simplifies our notation below.

DEFINITION: A symmetric strategy profile $\sigma$ is a *symmetric sequential equilibrium* (SSE) for $\Gamma(k_0)$ if for any $t \geq 0$ , history $h^{t-1} \in H^t$ with corresponding current capital $k_t$, and $x_t \in [0,1]$, the following conditions are satisfied:

(i) For any deviation strategy $\gamma \in \Sigma_G$ for the government,

$$\Phi_G(k_t, \sigma|_{(h^{t-1}, x_t)}) \geq \Phi_G(k_t, (\sigma_C|_{(h^{t-1}, x_t)}, \gamma|_{x_t})).$$

(ii) For any $\hat{\tau}_0 \in [\underline{\tau}, \overline{\tau}]^2$, let $(\hat{\tau}, \hat{q})$ be the stochastic tax policy-allocation generated by $\hat{\tau}_0$ and $\sigma|_{(h^{t-1}, x_t, \hat{\tau}_0)}$, so that taxes in period $t$ are $\hat{\tau}_0$ and taxes in subsequent periods are determined by $\sigma|_{(h^{t-1}, x_t, \hat{\tau}_0)}$. Then $\hat{q}$ is a competitive equilibrium of $\Gamma(k_t \mid \hat{\tau})$.

Let $\sigma$ be a strategy for $\Gamma(k_0, x_0)$. Then, $\sigma$ is an SSE for $\Gamma(k_0, x_0)$ if and only if when it is considered as a strategy for $\Gamma(k_0)$ with deterministic period 0 actions, $\sigma$ is an SSE of $\Gamma(k_0)$.

Conditions (i) and (ii), respectively, require that the government's and the households' continuation strategies be best responses (to each other) after any history $h^{t-1}$. A household's deviation at time $t-1$ is not detectable and cannot affect anybody else's future actions (including the government's). Thus, after such a deviation, the other households continue to play the competitive equilibrium of $\Gamma(k_t \mid \hat{\tau})$. This justifies our earlier assertion that a household need not specify its behavior following its own deviation. What the household does after a deviation is irrelevant: by Lemma 3, the household can never recoup the losses incurred at the time of the deviation.

## 5. SELF-GENERATION

The main idea of APS is that the sequential equilibria of an infinitely repeated game can be described recursively. A vector of lifetime payoffs is produced by a sequential equilibrium only if it is the weighted sum of current-period payoffs and a vector of lifetime payoffs produced by another sequential equilibrium (the continuation equilibrium). The set $V$ of sequential equilibrium payoffs then becomes the focus of the analysis.

In an $N$-person repeated game, the equilibrium value set $V$ is a subset of $\mathbf{R}^N$. Atkeson (1991) extends APS to dynamic games with a public state variable $k$. In an $N$-person dynamic game, the set of sequential equilibria depends on the initial state $k$ and the collection of equilibrium values is represented by an equilibrium value *correspondence* $V$, where $V(k) \subset \mathbf{R}^N$ is the set of all sequential equilibrium values when the initial state is $k$.

In our environment, the state variable is the distribution of privately held capital $k$. When the distribution of capital is degenerate and almost all households hold the same amount of capital $k$, a given sequential equilibrium delivers a lifetime utility $v$ to the government and to the measure one of households holding $k$ units of capital. It also delivers a lifetime utility $\hat{v}(\hat{k})$ for each household holding an off the equilibrium path level of capital $\hat{k} \neq k$ (and where $\hat{v}(k) = v$). A direct extension of APS would require an infinite dimensional equilibrium value correspondence $V$. For any $k$, $V(k)$ would represent the set of functions $\hat{v}(\hat{k})$ which can be produced by a sequential equilibrium. Such an infinite dimensional value correspondence would be difficult to characterize either computationally or analytically. Thus, a straightforward approach appears problematic.

In this section, we introduce a more efficient strategic dynamic programming method than the simplistic extension of APS discussed above. We show that it is sufficient to work with an equilibrium value correspondence $V$, where for each $k \in [\underline{k}, \overline{k}]$, $V(k)$ is a subset of $R^2$. For each SSE $\sigma$ of $\Gamma(k)$, one dimension of $V(k)$ represents the value of $\sigma$. The other dimension is the derivative at $\hat{k} = k$ of the function $\hat{v}(\hat{k})$ defined by $\sigma$. We call this object the marginal value of capital.

For any $\sigma \in \Sigma$ and $x_0 \in [0, 1]$, $\sigma|_{x_0}$'s and $\sigma$'s *marginal value of capital* are, respectively,

$$\Phi_C(k_0, \sigma|_{x_0}) := (1 - \tau_{k,0}(x_0)) f_k(k_0, \ell_0(x_0)) u_c(\ell_0(x_0), c_0(x_0))$$

$$\Phi_C(k_0, \sigma) := E_{x_0}\big[\Phi_C(k_0, \sigma|_{x_0})\big].$$

The former represents the increase of a household's utility had it started with an additional unit of capital and spent all the additional income on consumption in period 0. Together with the previously defined value of $\sigma$, we now denote $\Phi(k_0, \sigma) = (\Phi_C(k_0, \sigma), \Phi_G(k_0, \sigma))$.

DEFINITION: The *ex-ante and exp-post equilibrium value correspondences* are the set-valued functions $V : [\underline{k}, \overline{k}] \to \mathbf{R}^2$ and $V^R : [\underline{k}, \overline{k}] \to \mathbf{R}^2$ that to each initial capital per capita $k_0$ and any outcome $x_0$ associate the sets

$$V(k_0) := \{\, \Phi(k_0, \sigma) \mid \sigma \text{ is an SSE for } \Gamma(k_0) \,\}.$$

$$V^R(k_0) := \{\, \Phi(k_0, \sigma|_{x_0}) \mid \sigma \text{ is an SSE for } \Gamma(k_0) \,\}.$$

That is, $V(k_0)$ is the collection of all the pairs $(m, v)$ for which there exists an SSE $\sigma$ such that $m$ is its expected marginal value of capital for households in the first period and $v$ is the lifetime expected payoff for the government (and the representative household). $V^R(k_0)$ is the same as $V(k_0)$ but where $m$ and $v$ are calculated after the realization $x_0$ of $X_0$.

For every $k_0$ and $x_0, x_0' \in [0, 1]$, the subgames $\Gamma(k_0, x_0)$ and $\Gamma(k_0, x_0')$ are identical. Therefore it does not matter which outcome $x_0$ is selected to define $V^R(k_0)$, and $V^R(k_0)$ does not depend on $x_0$. That is, the *set* of possible payoffs after $x_0$ is realized does not depend on the actual realization of $x_0$. However, under a *particular* SSE, lifetime payoffs after the realization of $x_0$ can depend on $x_0$. For instance, an SSE $\sigma$ of $\Gamma(k_0)$ can select a (possibly different) SSE of $\Gamma(k_0, x_0)$ for each realization $x_0$. For instance, if $\sigma$ selects the SSE $\sigma_1$ of $\Gamma(k_0, x_0)$ when $x_0 \in [0, \lambda]$ and the SSE $\sigma_2$ when $x_0 \in (\lambda, 1]$, then

$$\Phi(k_0, \sigma) = \lambda \Phi(k_0, \sigma_1) + (1 - \lambda)\Phi(k_0, \sigma_2).$$

Hence, $V(k_0)$ is the convex hull of $V^R(k_0)$.

DEFINITION: A correspondence $W : [\underline{k}, \overline{k}] \to \mathbf{R}_+^2$ with compact and convex values is called a *value correspondence*. If $W$ is a value correspondence, for each $k \in [\underline{k}, \overline{k}]$, the upper and lower boundaries of $W(k)$ are represented by the functions

$$\overline{W}(k, m) := \max_v \{\, v \mid (m, v) \in W(k) \,\}$$

$$\underline{W}(k, m) := \min_v \{\, v \mid (m, v) \in W(k) \,\}.$$

By convention, $\overline{W}(k, m) = -\infty$ and $\underline{W}(k, m) = +\infty$ if there does not exist $v$ such that $(m, v) \in W(k)$. The lowest value in $W(k)$ is $\underline{v}_W(k) = \min_m \underline{W}(k, m)$.

We now derive a recursive characterization of $V$ and $V^R$. Let $\sigma$ be an SSE of $\Gamma(k_0)$ and $x_0 \in [0, 1]$. The strategy $\sigma|_{x_0}$ generates period zero path play $(\tau, \ell, c, k_+)$ and on-path continuation profile $\sigma|_{h_0}$, after history $h_0 = (x_0, \tau)$, with value $(m_+, v_+) = \Phi(k_+, \sigma|_{h_0})$. Since $\sigma|_{x_0}$ is an SSE of $\Gamma(k_0, x_0)$, $\sigma|_{h_0}$ is an SSE of $\Gamma(k_+)$, and therefore $(m_+, v_+) \in V(k_+)$. Let $\xi = (\tau, \ell, c, k_+, m_+, v_+) \in R^7$ and $W$ be an arbitrary value correspondence. When $W = V$, the following definitions of *consistency* and *admissibility* capture all the conditions for $(\tau, \ell, c, k_+)$ to be the period 0 outcome of an SSE $\sigma$ and for $(m_+, v_+) = \Phi(k_+, \sigma|_{h^0})$.

DEFINITION: A vector $\xi = (\tau, \ell, c, k_+, m_+, v_+)$ is *consistent* with respect to the value correspondence $W$ at endowment $k \in [\underline{k}, \overline{k}]$ if $(\ell, c, k_+) \in CE(k, \tau, m_+)$, $k_+ \in [\underline{k}, \overline{k}]$, and $(m_+, v_+) \in W(k_+)$. The vector $\xi$'s value and marginal value of capital are[6]

$$\Psi_G(k, \xi) := (1 - \beta) \left[ u(\ell, c) + g(\tau_k f_k(k, \ell)k + \tau_\ell f_\ell(k, \ell)\ell) \right] + \beta v_+$$
$$\Psi_C(k, \xi) := (1 - \tau_k) f_k(k, \ell) u_c(\ell, c).$$

For each $\tau' \in [\underline{\tau}, \overline{\tau}]^2$, let

$$\pi_W(k, \tau') := \min_{(\ell', c', k'_+, m'_+, v'_+)} \Psi_G(k, \xi')$$
$$\text{s.t.} \qquad \xi' = (\tau', \ell', c', k'_+, m'_+, v'_+)$$
$$\text{is consistent w.r.t. } W \text{ at } k,$$

and define

$$\overline{\pi}_W(k) := \max_{\tau' \in [\underline{\tau}, \overline{\tau}]^2} \pi_W(k, \tau').$$

The vector $\xi$ is *admissible* with respect to $W$ at endowment $k$ if it is consistent with respect to $W$ at $k$ and

$$\Psi_G(k, \xi) \geq \overline{\pi}_W(k).$$

Admissibility is the government's incentive constraint. When the government announces unexpected tax rates $\tau'$, the households' beliefs are manipulated in the ensuing subgame so as to yield the worst possible payoff for the government. Suppose this entails the belief that the marginal utility of investment will be $m'_+$, and that given this belief, each household saves $k'_+$ for production next period. Since the choice of continuation value $v'_+$ for next period does not affect the optimization problem of the households this period (in particular, it does not affect the choice $k'_+$ of capital investment), the worst punishment must choose the continuation value $v'_+ = \underline{W}(k'_+, m'_+)$.[7] Thus, as in Abreu's (1986,

---

[6]  For a strategy profile $\sigma$ of $\Gamma(k)$, we previously defined its value $\Phi_G(k, \sigma)$ and marginal value of capital $\Phi_C(k, \sigma)$. Although $\Phi$ and $\Psi$ are similar functions, they are defined on different domains.

[7]  The only reason to require in the definition of admissibility that the correspondence $W$ be compact-valued is to guarantee that in the definition of $\pi_W(k, \tau')$, the minimum is attained. We could define admissibility for an arbitrary correspondence $W$ if we replace this minimum with an infimum.

1988) optimal punishment for firms in a cartel, without loss of generality, the definition of $\pi_W(k, \tau')$ considers extreme punishments only.

DEFINITION: For a value correspondence $W$, let

$$B(W)(k) = co\big(\{\, \Psi(k, \xi) \mid \xi \text{ is admissible w.r.t. } W \text{ at } k \,\}\big)$$

where $co(X)$ refers to the convex hull of $X \subset \mathbf{R}^2$. We say a correspondence $W$ is *self-generating* if $W(k) \subset B(W)(k)$ for all $k$ (which we write simply as $W \subset B(W)$).

For an arbitrary value correspondence $W$ and $k$, the computation of $B(W)(k)$ can be divided into two steps. Let

$$\hat{B}(W)(k) := \{\, \Psi(k, \xi) \mid \xi \text{ is consistent w.r.t. } W \text{ at } k \,\}.$$

Then,

$$B(W)(k) = co\big(\{\, (m, v) \in \hat{B}(W)(k) \mid v \geq \overline{\pi}_W(k) \,\}\big).$$

Thus, consistency and the government's incentive constraints are "separable": the latter is only required to truncate, for each $k$, the set of consistent vector values obtained when the constraint is neglected. The previous observation implies that $\underline{v}_{B(W)}(k) = \min \{\, v \mid (m, v) \in B(W)(k) \,\} = \overline{\pi}_W(k)$ for all $k$. This observation is used in the proof of Theorem 5.

The following theorem is an adaptation of the similar result proven in APS for repeated games. For completeness, its proof and the proof of Theorem 2 below are presented in the Appendix.

THEOREM 1: *(Self-Generation) Let $W$ be a value correspondence. If $W$ is self-generating (or $W \subset B(W)$), then $B(W) \subset V$.*

THEOREM 2: *If $W$ is an upper semicontinuous (usc) value correspondence, then $B(W)$ is a usc value correspondence.*

LEMMA 4: *graph(V) is a bounded set.*

PROOF: Let $\sigma$ be any SSE of $\Gamma(k_0)$. The government's income is always nonnegative, and a household can choose not to work and not to save in every period, and thus guarantee itself a total payoff of at least $\underline{v} = u(0, 0) + g(0)$ in every period. On the other hand, a household cannot get more than $\overline{v} = u(0, \overline{k}) + g(\overline{k})$ in any period. Hence $\underline{v} \leq \Phi_G(k_0, \sigma) \leq \overline{v}$.

Now, $\Phi_C(k_0, \sigma) \geq 0$ because for each history $h^{t-1}$ and outcome $x_t$, $\tau_{k,t} \leq \overline{\tau} < 1$, the marginal productivity of capital is positive, and the marginal utility of consumption is positive. Finally, Lemma 2 establishes the bound $\Phi_C(k_0, \sigma) \leq \overline{m}$. Therefore, graph$(V) \subset [\underline{k}, \overline{k}] \times [0, \overline{m}] \times [\underline{v}, \overline{v}]$, and graph$(V)$ is bounded. ∎

Similar to APS's main result, the equilibrium correspondence $V$ is the largest fixed point of the map $B$. That is, $V = B(V)$, and $W \subset V$ for any self-generating correspondence $W$ with compact and convex values.

THEOREM 3: *(Factorization) The equilibrium correspondence $V$ is the largest value correspondence $W$ such that $W = B(W)$. Moreover, $V$ is usc.*

PROOF: Let $\mathrm{cl}(V)$ denote the correspondence whose graph is the closure of $\mathrm{graph}(V)$. Since $\mathrm{graph}(V)$ is bounded, $\mathrm{graph}(\mathrm{cl}(V))$ is compact. A correspondence with a compact graph is usc. Therefore, $\mathrm{cl}(V)$ is usc. Since $V(k)$ is convex, $\mathrm{cl}(V)(k)$ is also convex for all $k \in [\underline{k}, \overline{k}]$.

Pick any SSE $\sigma$ of $\Gamma(k_0)$, and let $(m, v) := \Phi(k_0, \sigma)$. For each $x \in [0, 1]$, let $\tau(x) = \sigma_G(0)(x)$ and $(\ell(x), c(x), k_+(x))$ be the allocation produced by $\sigma_C(0)(x, \tau(x))$. That is, $(x, \tau(x), \ell(x), c(x), k_+(x))$ is the equilibrium outcome in the first period (when the realization of the random public device is $x$). Also define $(m_+(x), v_+(x)) = \Phi(k_+(x), \sigma|_{h_0})$, where $h_0 = (x, \tau(x))$. Then, the vector $\xi(x) = (\tau(x), \ell(x), c(x), k_+(x), m_+(x), v_+(x))$ is admissible w.r.t. $V$ at $k_0$, and

$$(m, v) = \int_0^1 \Psi(k, \xi(x))\, dx.$$

This implies that $V \subset B(\mathrm{cl}(V))$ (since $\mathrm{cl}(V)$ is a value correspondence, $B(\mathrm{cl}(V))$ is well defined). By Theorem 2, $\mathrm{graph}(B(\mathrm{cl}(V)))$ is compact, and therefore $\mathrm{cl}(V) \subset B(\mathrm{cl}(V))$. That is, $\mathrm{cl}(V)$ is self-generating, and by Theorem 1, we must have that $B(\mathrm{cl}(V)) \subset V$. Hence $\mathrm{cl}(V) \subset V$, and $V$ has a closed graph. That is, $V = \mathrm{cl}(V)$. Moreover, the previous inclusions imply that $V \subset B(V) \subset V$, so $V = B(V)$. Finally, if a value correspondence $W$ satisfies $W \subset B(W)$, Theorem 1 imples that $W \subset V$. ∎

The previous theorem establishes that for each $k \in [\underline{k}, \overline{k}]$ and $m$ in the appropriate range, there exists an (worst) equilibrium $\sigma$ for $\Gamma(k)$ such that $\Phi_G(k, \sigma) = \underline{V}(k, m)$, and thus the extreme punishments invoked in the definition of admissibility *can* be delivered in equilibrium.

We finally extend the algorithm introduced by APS for repeated games to compute the equilibrium value correspondence. It is easy to see that $B$ is monotone. That is, if $W$ and $W'$ are two value correspondences such that $W \subset W'$, then $B(W) \subset B(W')$. Suppose that $W_0$ is a value correspondence, such that $V \subset W_0$ and $W_0 \supset B(W_0)$. Construct the sequence of value correspondences $\{W_n\}_{n=0}^{\infty}$ inductively as follows:

$$W_{n+1} = B(W_n) \quad \text{for all } n \geq 0.$$

Then $W_n \supset W_{n+1} \supset V$ for all $n \geq 0$. Since the sequence is decreasing (in the sense of set inclusion), it has a limit $W_{\infty}$. That is, $W_{\infty}$ is the value correspondence defined by

$$W_{\infty}(k) = \lim_{n \to \infty} W_n(k) = \bigcap_{n \geq 0} W_n(k)$$

for each $k \in [\underline{k}, \overline{k}]$. By a simple limit argument we can establish the following result.

THEOREM 4: $W_{\infty} = V$.

We use this algorithm to compute the solution to an example in Section 7.

## 6. Best Equilibria and Steady States

We now study general properties of equilibria. In particular, we focus on best equilibria (given an initial capital stock $k$ and marginal value of capital $m$) and on long run behavior. As in Chamley (1986) and Judd (1985), we study the steady state of best equilibria and compare it with the steady state of the optimal policy with commitment. Although our model is simpler than that of Chamley, the steady state of the optimal tax policy with commitment (if it exists) exhibits the same properties: the tax on capital is 0. Without commitment, this is not always the case. Nevertheless, we obtain a sharp characterization of the steady state of the best equilibrium when its corresponding tax rate on capital is positive. First, the steady state delivers the *best* equilibrium payoff associated with the steady state level of capital $k^s$ and the steady state marginal value of capital $m^s$. That is, its vector value is on the upper boundary of the set $V(k^s)$. Second, the steady state delivers the *worst* equilibrium payoff associated with the level of capital $k^s$. So its value is also on the lower boundary of $V(k^s)$.

We say that an SSE $\sigma$ of $\Gamma(k_0)$ is a *best* SSE if its vector value $\Phi(k_0, \sigma)$ is on the upper boundary of $V(k_0)$. The next lemma establishes that if $\sigma$ is a best SSE, then all its equilibrium continuation strategies are best SSE's. That is, for any equilibrium path history $h^{t-1}$, the vector value of $\sigma|_{h^{t-1}}$ is on the upper boundary of $V(k_t)$ (where $k_t$ is the corresponding capital stock associated with $h^{t-1}$).

LEMMA 5: *If the vector $\xi = (\tau, \ell, c, k_+, m_+, v_+)$ is admissible with respect to $V$ at $k$ and $\Psi_G(k, \xi) = \overline{V}(k, m)$ for some $m$, then $v_+ = \overline{V}(k_+, m_+)$.*

PROOF: Suppose not. Choose $\hat{v}_+$ such that $v_+ < \hat{v}_+ \leq \overline{V}(k_+, m_+)$. Since the government's incentives are strengthened when its continuation value is increased, $\hat{\xi} = (\tau, \ell, c, k_+, m_+, \hat{v}_+)$ is also admissible w.r.t. $V$ at $k$, and $\Psi_G(k, \hat{\xi}) > \Psi_G(k, \xi)$. This contradicts $\Psi_G(k, \xi) = \overline{V}(k, m)$. ∎

For completeness, we briefly discuss the characteristics of a worst equilibrium. In every period, there are two instances where it is relevant to construct worst equilibria: at the beginning of the period (before the outcome of the random device is observed) and after the government chooses tax rates. By symmetry, we say that $\sigma$ is a worst SSE of $\Gamma(k_0)$ if $\Phi(k_0, \sigma)$ is on the lower boundary of $V(k_0)$. We discussed the nature of a worst equilibrium (or punishment) after the government moves when we defined $\overline{\pi}_V(k_0)$. A worst punishment at that juncture involves manipulating the beliefs of the households (the value of $m_1$) so that they coordinate in a relatively bad outcome for the current period and capital stock $k_1$ for the next period. The continuation of such a worst punishment is indeed a worst SSE of $\Gamma(k_1)$ with value $\underline{V}(k_1, m_1)$.

However, in general, the continuation of a worst SSE of $\Gamma(k_0)$ is not a worst SSE of $\Gamma(k_1)$. This holds because in a worst equilibrium (one that attains $\underline{V}(k_0, m_0)$), the incentive constraint on the government typically binds. That is, if $\xi = (\tau, \ell, c, k_1, m_1, v_1)$ is an admissible vector w.r.t. $V$ at $k_0$ representing a worst SSE of $\Gamma(k_0)$ and $m_0 = \Psi_C(k_0, \xi)$, then $\Psi_G(k_0, \xi) = \underline{V}(k_0, m_0) = \overline{\pi}_V(k_0)$. Therefore, usually, there exist consistent vectors $\hat{\xi} = (\tau, \ell, c, k_1, m_1, \hat{v}_1)$ with $\underline{V}(k_1, m_1) \leq \hat{v}_1 < v_1$. But, since lowering continuation utility $v_1$ lowers present utility, these consistent vectors violate the government incentive

constraint and thus cannot be used to lower the government's utility even further.

We say that $\xi^s = (\tau^s, \ell^s, c^s, k^s, m^s, v^s)$ is a *steady state vector* if

$$(\ell^s, c^s, k^s) \in CE(k^s, \tau^s, m^s) \quad \text{and} \quad (m^s, v^s) = \Psi(k^s, \xi^s).$$

That is, $\xi^s$ is a steady state vector if households choosing $(\ell^s, c^s, k^s)$ in every period is a competitive equilibrium of $\Gamma(k^s \mid \{\tau^s\})$ (the economy starting from $k^s$ with taxes equal to $\tau^s$ at all dates). If a steady state vector $\xi^s$ is also admissible (i.e., the government's incentive constraint is satisfied), then $\xi^s$ describes the constant path of play of an SSE of $\Gamma(k^s)$ (or $\xi^s$ is an *equilibrium steady state vector*).

The next theorem characterizes the steady state of a best equilibrium. We show that if a best equilibrium reaches a steady state with corresponding vector $\xi^s$, then in the steady state, either capital taxes are zero or the steady state represents a *worst* equilibrium for $\Gamma(k^s)$.

THEOREM 5: *Let $\sigma$ be an SSE for $\Gamma(k_0)$. Suppose that $\Phi_G(k_0, \sigma) = \overline{V}(k_0, \Phi_C(k_0, \sigma))$ and that $\sigma$ converges to the steady state vector $\xi^s = (\tau^s, \ell^s, c^s, k^s, m^s, v^s)$. Then either (1) $\tau_k^s = 0$ or (2) $v^s = \underline{v}_V(k^s)$.*

Before proceeding to the proof, we note some implications. Since $\Phi(k_0, \sigma)$ is on the upper boundary of $V(k_0)$, along the outcome path, $\Phi(k_t, \sigma|_{h^{t-1}})$ is on the upper boundary of $V(k_t)$, for the corresponding capital stock $k_t$, for all $t \geq 0$. Thus, by continuity, $(m^s, v^s)$ is on the upper boundary of $V(k^s)$. That is, $v^s = \overline{V}(k^s, m^s)$. When $\tau_k^s > 0$, the theorem asserts that $v^s = \underline{v}_V(k)$. Therefore, if $\tau_k^s > 0$, $m^s$ must be either the lowest or highest marginal value of capital possible in $V(k^s)$ and $(m^s, v^s)$ is at the intersection of the upper and lower boundaries of $V(k^s)$.

PROOF: To simplify the notation, we write the proof for the game $\Gamma(k_0)$ without public randomization. Alternatively, assume that the equilibrium value correspondence of the game without public randomization is convex valued (and thus the randomization device is superfluous). For any (deterministic) outcome path $q = \{(\tau_{k,t}, \tau_{\ell,t}, \ell_t, c_t, k_{t+1})\}_{t=0}^{\infty}$ and any $t \geq 0$, let

$$S_t(q) = (1 - \beta) \sum_{i=0}^{\infty} \beta^i [u(\ell_{t+i}, c_{t+i}) + g(G_{t+i})],$$

where $G_t = \tau_{k,t} f_k(k_t, \ell_t) k_t + \tau_{\ell,t} f_\ell(k_t, \ell_t) \ell_t$. Following Chamley, we can view the government choices of taxes to be equivalent to the choices of after-tax wage $w_t$ and "interest rate" on capital $r_t$, where

$$1 + r_t := (1 - \tau_{k,t}) f_k(k_t, \ell_t) \quad \text{and} \quad w_t := (1 - \tau_{\ell,t}) f_\ell(k_t, \ell_t).$$

We also use the standard notation of $u_{\ell,t}$ and $u_{c,t}$ for $u_\ell(\ell_t, c_t)$ and $u_c(\ell_t, c_t)$, respectively. Then if $\Phi_G(k_0, \sigma) = \overline{V}(k_0, \Phi_C(k_0, \sigma))$,

$$\Phi_G(k_0, \sigma) = \max_q \quad S_0(q)$$

$$\text{s.t.} \quad (1 + r_0) u_{c,0} = \Phi_C(k_0, \sigma)$$
$$(\ell_t, c_t, k_{t+1}) \in CE(k_t, \tau_t(w_t, r_t, k_t, \ell_t), (1 + r_{t+1}) u_{c,t+1})$$
$$c_t + k_{t+1} + G_t \leq f(k_t, \ell_t)$$
$$S_t(q) \geq \overline{\pi}_V(k_t),$$

15

where $\tau_t(w_t, r_t, k_t, \ell_t)$ refers to the tax pair implied by $r_t$ and $w_t$ given $k_t$ and $\ell_t$. The first constraint requires that the initial marginal value of capital be delivered. The second requires that the trajectory of labor, consumption and savings constitutes a competitive equilibrium of the dynamic economy that arises when the government fixes its tax policy at $\tau$. The third constraint is the resource constraint. The final is the government's incentive constraint. It says that what the government expects in equilibrium from period $t$ onward has to be no less than the worst punishment value for the game $\Gamma(k_t)$.

Suppose that the solution converges to the steady state represented by the vector $\xi^s = (\tau_k^s, \tau_\ell^s, \ell^s, c^s, k^s, m^s, v^s)$. Then $k^s \geq \underline{k} > 0$. It is easy to see that $\ell^s > 0$; otherwise, $f(k^s, \ell^s) = 0$, which is incompatible with the government's incentive constraint. Since $k^s > 0$ and $\ell^s > 0$, we must also have that $c^s > 0$ (otherwise the household could increase its utility by decreasing its savings). If $\ell^s < 1$ (so that the household's choices are "interior"), the competitive equilibrium requirement for $t$ sufficiently large is equivalent to the first order conditions

$$u_{\ell,t} + w_t u_{c,t} = 0, \quad u_{c,t} = \beta(1 + r_{t+1})u_{c,t+1}, \quad \text{and} \quad c_t + k_{t+1} = (1 + r_t)k_t + w_t\ell_t.$$

For simplicity, assume that the household's trajectory is "interior" (the alternative case, when $\ell^s = 1$ can be dealt with in a similar fashion). Then, we can replace the competitive equilibrium requirement with its corresponding first-order conditions to obtain

$$\Phi_G(k_0, \sigma) = \max_q \quad S_0(q)$$

$$\text{s.t.} \quad (1 + r_0)u_{c,0} = \Phi_C(k_0, \sigma)$$
$$u_{\ell,t} + w_t u_{c,t} = 0$$
$$u_{c,t} = \beta(1 + r_{t+1})u_{c,t+1}$$
$$c_t + k_{t+1} = (1 + r_t)k_t + w_t\ell_t$$
$$c_t + k_{t+1} + G_t \leq f(k_t, \ell_t)$$
$$S_t(q) \geq \overline{\pi}_V(k_t).$$

Let $\lambda_t\beta^t$, $\mu_t\beta^t$, and $\gamma_t\beta^t$ be the Lagrange multipliers corresponding to the last three constraints (which must hold for all $t$). It is easy to see that the resource constraint must be active for all $t \geq 0$. If the government's incentive constraint in period $t$ is not active, then $\gamma_t = 0$ and the optimality condition associated with $k_t$ is

$$0 = \lambda_{t-1} - \lambda_t\beta(1 + r_t) + \mu_{t-1} - \mu_t\beta f_{k,t}.$$

Assume first that the government's incentive constraint is not active in the limit, that is, that

$$u(\ell^s, c^s) + g(G^s) > \overline{\pi}_V(k^s).$$

Then, by continuity, $S_t(q) > \overline{\pi}_V(k_t)$ for all $t$ sufficiently large. In the limit, the second constraint implies that $\beta(1 + r^s) = 1$. Therefore, in the limit, the optimality condition for $k_t$ implies that $1 - \beta f_k(k^s, \ell^s) = 0$. The last two equalities imply that $1 + r^s = f_k(k^s, \ell^s)$, or that $\tau_k^s = 0$. This is the Chamley result (and proof).[8]

---

Recall from Section 5 that for all $k$ and $W$, $\underline{v}_{B(W)}(k) = \overline{\pi}_W(k)$. Hence, since $V = B(V)$, $\overline{\pi}_V(k^s) = \underline{v}_V(k^s) = \min_m \underline{V}(k^s, m)$. Thus, if the government's incentive constraint is active in the limit, we have

$$\lim_{t \to \infty} \Phi(k_{t+1}, \sigma|_{h^t}) = \lim_{t \to \infty} S_t(q) = u(\ell^s, c^s) + g(G^s) = \min_m \underline{V}(k^s, m). \qquad \blacksquare$$

Whether Chamley's result attains without commitment depends on the discount factor $\beta$ (and on $\overline{\tau}$ as well). In equilibrium, the government's temptation to deviate to a higher capital tax and lower labor tax in the first period, for example, is deterred by a lower continuation value for the government. But, if the discount factor is low, the former effect dominates, and the government cannot resist the temptation. In this case, the steady state associated with zero capital taxes is not sustainable.

For each $k \in [\underline{k}, \overline{k}]$, let $L(k)$ denote the locus of vector values $\Psi(k, \xi)$ for steady state vectors $\xi = (\tau, \ell, c, k, m, v)$. That is, $L(k)$ represents the vector values of all steady states that maintain the same level of capital stock $k$. For some $k$'s, the collection of steady states and $L(k)$ is empty. But, other $k$'s admit multiple steady state vectors and $L(k)$ is a curve in $\mathbf{R}^2$.

In our computed examples (see Section 7 below), when the best SSE converges to a steady state with capital $k^s$, where the government's constraint is binding (so Chamley's result does not attain), $V(k^s)$ and $L(k^s)$ are tangent, and the point of tangency corresponds to the highest point in $L(k^s)$ (that is, the point with the highest value for the government). From Theorem 5 then this tangency point is at the intersection of the upper and lower boundaries of $V(k^s)$ and the cusp of $L(k^s)$. The next lemma establishes that for an arbitrary $k$, if $V(k)$ and $L(k)$ are tangent, then the highest point in $L(k)$ must be a tangency point. Hence, if the tangency point between $V(k)$ and $L(k)$ is unique, it must coincide with the highest point in $L(k)$.

LEMMA 6: *Let* $k \in [\underline{k}, \overline{k}]$ *and* $(\hat{m}, \hat{v})$ *be such that* $\hat{v} = \max \{ v \mid (m, v) \in L(k) \}$. *Suppose that* $V(k) \cap L(k) \neq \emptyset$. *Then* $(\hat{m}, \hat{v}) \in V(k)$.

PROOF: Let $\hat{\xi} = (\hat{\tau}, \hat{\ell}, \hat{c}, k, \hat{m}, \hat{v})$ and $\xi = (\tau, \ell, c, k, m, v)$ be two steady state vectors such that $\Psi(\xi) = (m, v) \in V(k)$. Then $\Psi(k, \hat{\xi}) = (\hat{m}, \hat{v})$ and $\hat{v} \geq v \geq \underline{v}_V(k)$. Therefore $\hat{\xi}$ represents the stationary outcome path of an SSE of $\Gamma(k)$ and $(\hat{m}, \hat{v}) \in V(k)$. $\qquad \blacksquare$

For each $k$, let $\hat{v}(k)$ denote the highest value for the government in $L(k)$. The previous lemma implies that $V(k) \cap L(k) \neq \emptyset$ for all $k$ such that $\hat{v}(k) \geq \underline{v}_V(k)$. Our computational results also show that when the government's incentive constraint is binding in the limit steady state of the best SSE, then the corresponding limit capital stock is the largest $k$ such that $\hat{v}(k) \geq \underline{v}_V(k)$.

## 7. AN EXAMPLE

In this section we present an example. We have implemented the algorithm described in Section 5 (Theorem 4) in a computer program. We use this program to compute numerically the equilibrium value correspondence of the example. We have also implemented a numerical algorithm to compute the steady states and corresponding vector values for

each capital stock. We are thus able to identify the Chamley-Judd (full commitment with distortionary taxes) and no-commitment steady state outcomes.

In the example we have made the following choices:

$$u(\ell, c) = 2\sqrt{c} - (1 - \ell^{1.1})^{.9}, \quad g(G) = 2\sqrt{G}, \quad \text{and} \quad \beta = 0.9,$$
$$f(k, \ell) = k^{.3}\ell^{.7}, \quad \text{and} \quad \delta = 0.05,$$

where $\delta$ represents the capital depreciation rate. As we explained in footnote 2, we make the assumption of full depreciation (i.e., $\delta = 1$) throughout the paper for convenience only. Our results also apply to the case in which $\delta < 1$.[9] To reduce the size of the numerical problem, we assume that values for $\tau_k$ and $\tau_\ell$ are restricted to be in the discrete grid $\{0, .1, .2, .25, .26, \ldots, .29, .3, .4, \ldots, .9\}$. This can be viewed as an institutional constraint, where the government is not allowed other intermediate values of taxes. From .25 to .30 we made the grid finer to allow a closer approximation of the steady state.

The equation $k = (1 - \delta)k + f(k, 1)$ implies $\overline{k} = 20^{10/7}$. But this upper bound is too generous. We set instead $\overline{k} = 1.57$, which is the largest capital stock that can be sustained in a steady state with full commitment (when $\tau_{k,t} = \tau_{\ell,t} = 0$ for all $t \geq 0$). Similarly, we set $\underline{k} = 0.0044$, which is the smallest capital stock sustained in a steady state with full commitment (when $\tau_{k,t} = \tau_{\ell,t} = .9$ for all $t \geq 0$). We choose a grid for $k$ with 315 points uniformly spaced between $\underline{k}$ and $\overline{k}$ (or approximately every .005 units).
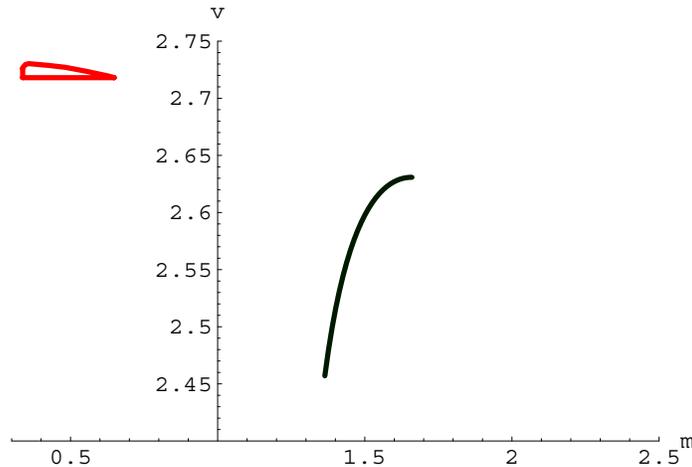


**Figure 1: Equilibrium Value Correspondence and Steady States for $k = 1.178$**

Figure 1 displays $V(1.178)$ (the set in the upper left of the figure) and the locus $L(1.178)$ (the arc in the lower right of the figure). The value $k = 1.178$ is the capital in the steady state of the solution to the Ramsey problem with commitment – the "Chamley-Judd" steady state. The uppermost point of the arc is obtained when $\tau_k = 0$ and $\tau_\ell = .28$. As can be seen in the figure, the Chamley-Judd steady state cannot be supported as an

---

[9] Since investments are irreversible, the household's problem must now include the constraint $k_{t+1} \geq (1 - \delta)k_t$ for all $t \geq 0$. Also, the definition of the marginal value of capital must be changed to $m_t = E[\sum_{s=t}^{\infty}(\beta(1 - \delta))^{s-t}(1 - \tau_{k,s})f_k(k_s, \ell_s)u_c(\ell_s, c_s)]$ or, recursively, $m_t = E[(1 - \tau_{k,t})f_k(k_t, \ell_t)u_c(\ell_t, c_t) + \beta(1 - \delta)m_{t+1}]$.

SSE. In particular, this implies that for any initial capital, the solution to the Ramsey problem with commitment cannot be supported as an SSE for the parameters we have chosen.

Figure 2 displays, as a function of the initial capital $k$, the value $\hat{v}(k)$ of the best steady state and the value $\underline{v}_V(k)$ of the worst equilibrium. At $k = .743$ these functions intersect. For $k > .743$ the worst equilibrium has a greater value than the best steady state and thus these $k$ values cannot be sustained as a steady state.[10] (This is likewise true for $k < .065$.)
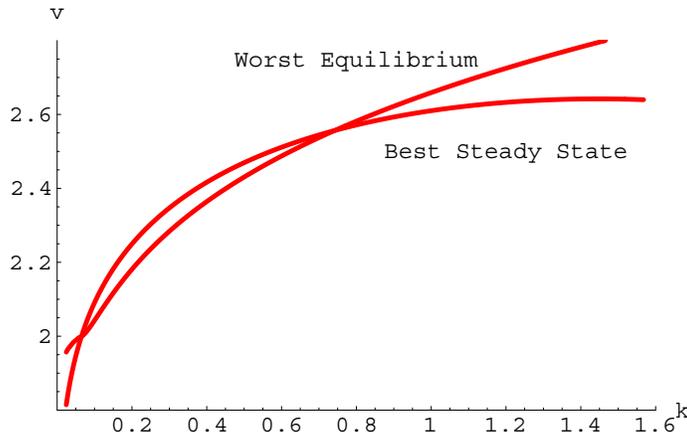


**Figure 2: Value of Best Steady State and Worst Equilibrium**
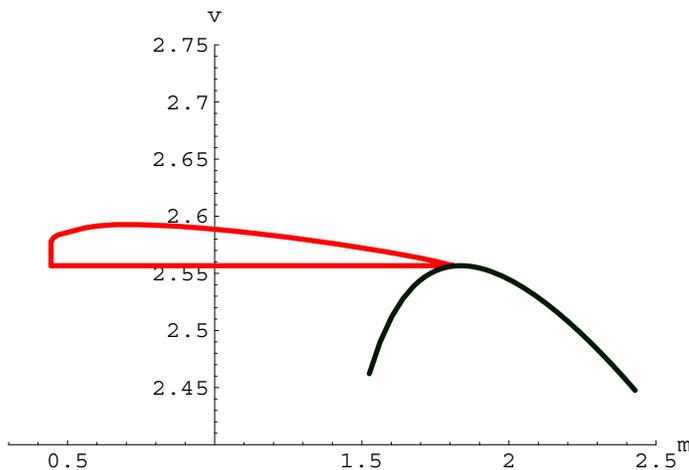


**Figure 3: Equilibrium Value Correspondence and Steady States for $k = .743$**

Figure 3 displays $V(0.743)$ and $L(0.743)$ (the arc touching the lower right of the set). The leftmost point of $L(0.743)$ attains when $\tau_{k,t} = 0.38$ and $\tau_{\ell,t} = 0$ for all $t \geq 0$,

---

[10]    Benhabib and Rustichini (1997) conjecture that allowing capital subsidies would allow the sustainability of steady state capital levels greater than the Chamley-Judd level even if the Chamley-Judd level is not sustainable as a steady state. For these parameters and all others we examined, allowing capital subsidies does not alter the relationship between these two curves. In particular, the curves do not intersect again at $k > 1.178$ as would be necessary for their conjecture to hold in our environment.

and the rightmost point attains when $\tau_{k,t} = 0$ and $\tau_{\ell,t} = 0.54$ for all $t \geq 0$. The only point of intersection between $V(0.743)$ and $L(0.743)$ represents the steady state of the best equilibrium without commitment (from any initial capital) and is attained when $\tau_{k,t} = 0.27$ and $\tau_{\ell,t} = 0.26$ for all $t \geq 0$. As stated in Theorem 5, this intersection coincides with the right corner of $V(0.743)$, where the upper and lower boundaries of $V(0.743)$ intersect, and the peak of the locus $L(0.743)$.

Figure 4 displays the time path of capital and Figure 5 the time path of $\tau_k$ for three different initial capital stocks, $k_0 = .054$, $k_0 = .743$, and $k_0 = 1.178$. (The last two are the steady states of the no-commitment and full commitment environments). The three paths converge to the same level of capital stock $(.743)$ and the same capital tax $(.27)$. In the three paths, the capital tax is initially maximal. From the low initial capital, capital taxes start at the maximal value, become lower than their steady state value (presumably to cause faster capital accumulation) and then converge (subject to small grid induced fluctuations) to their steady state value. From higher capital stocks, capital taxes are maximal longer and the steady state capital stock is approached from below.
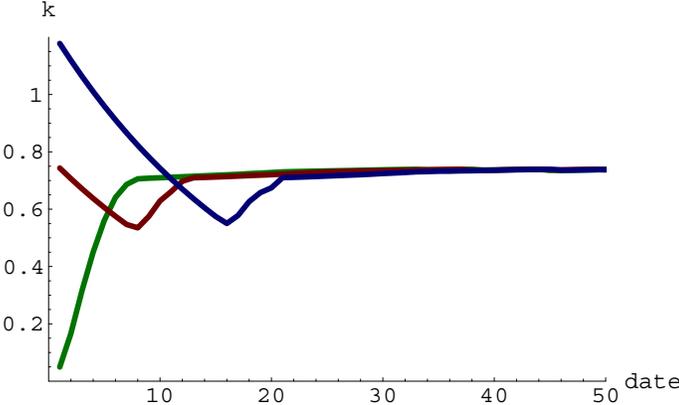


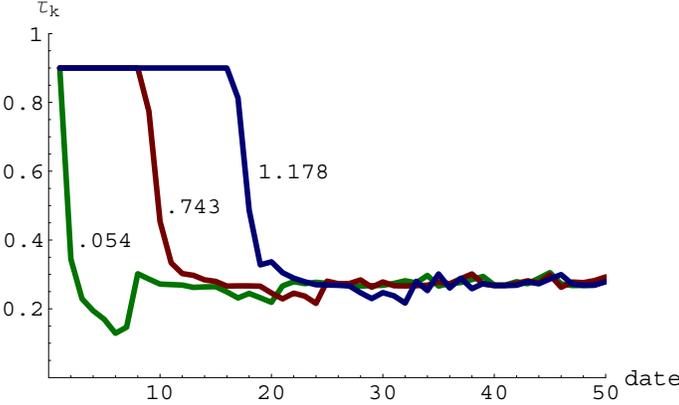Figure 4: Time Paths of Capital from $k = \{.054, .743, 1.178\}$



Figure 5: Time Paths of Capital Taxes from $k = \{.054, .743, 1.178\}$

20

## 8. Conclusion

We have elaborated a method for studying dynamic games with anonymous players. This class of games includes many standard "policy games", where the government chooses a policy (taxes, debt, or inflation) and there is a continuum of agents (households). The method uses the idea that since each household's actions are unobservable and do not affect prices, the household's intertemporal incentives can be captured completely by a single parameter (the marginal value of capital in our model). Although we have developed these ideas in the context of a dynamic game (with the capital stock as a state variable), a similar method would apply to a simpler class of infinitely repeated anonymous games.

The high complexity of these games often makes it impossible to find analytic solutions. However, the extension of the algorithm proposed by APS produces a sequence of approximations to the equilibrium correspondence that converges monotonically and provides a practical method for constructing the equilibrium correspondence.

## 9. Appendix

PROOF OF LEMMA 1: If the hypothesis is false, then for all $\hat{k}_0 > 0$, there exists $k_0 \geq \hat{k}_0$, $x^{t-1} \in [0,1]^t$, a stochastic tax policy $\tau$, and a period $t$ such that $k_t(x^{t-1}) < \hat{k}_0$ in some competitive equilibrium of $\Gamma(k_0 \mid \tau)$. Since the continuation of a competitive equilibrium is itself a competitive equilibrium, without loss of generality, we can assume that $t = 1$. Thus, by contradiction, let us assume that the decreasing sequence $\{\hat{k}_0^n\}_{n=0}^{\infty}$, with $\lim_{n \to \infty} \hat{k}_0^n = 0$, is such that for each $n$ there exist $k_0 > \hat{k}_0^n$, tax policy $\tau^n$, $\{x_t^n\}_{t=0}^{\infty}$, and competitive equilibrium $\{(\ell_t^n, c_t^n, k_{t+1}^n)\}_{t=0}^{\infty}$ of $\Gamma(k_0 \mid \tau^n)$ for which $k_1^n(x_0^n) < \hat{k}_0^n$.

For the rest of the proof we omit the random outcomes. Thus, for example, we write $k_1^n$ instead of $k_1^n(x_0^n)$ and $\ell^n$ instead of $\ell_1^n(x_0^n, x_1^n)$.

CLAIM 1: For $t = 0$ or $t = 1$, there does not exist a subsequence $\{n_r\}$ such that $\lim_{r \to \infty} \ell_t^{n_r} = 0$ and $\lim_{r \to \infty} \ell_t^{n_r} / k_t^{n_r} < +\infty$.

Proof of Claim 1: By contradiction, suppose that subsequence $\{n_r\}$ is such that $\lim_{r \to \infty} \ell_t^{n_r} = 0$ and $\lim_{r \to \infty} \ell_t^{n_r} / k_t^{n_r} = R < +\infty$. Since $f_\ell$ is homogeneous of degree 0, we have that

$$f_\ell(k_t^{n_r}, \ell_t^{n_r}) = f_\ell(1, \ell_t^{n_r} / k_t^{n_r}) \to f_\ell(1, R) > 0.$$

Labor taxes are uniformly bounded above by $\overline{\tau} < 1$, so after-tax wages in period $t$ are bounded below by $(1 - \overline{\tau}) f_\ell(1, R)/2 > 0$, say, for $n_r$ sufficiently large. By (A2), $\ell_t^{n_r} \to 0$ implies that $u_\ell(\ell_t^{n_r}, c_t^{n_r}) \to 0$. Therefore, for any positive constant $\gamma$, $-u_\ell(\ell_t^{n_r}, c_t^{n_r}) < \gamma(1 - \overline{\tau}) f_\ell(1, R)/2$, for $n_r$ sufficiently large. Let

$$\gamma := \inf_{\ell \in [0,1)} u_c(\ell, \overline{k})$$

(where $\overline{k} = f(\overline{k}, 1)$), which is positive by assumption. Suppose the household works a little harder in period $t$ and immediately consumes the additional income. Then, its total utility increases by

$$u_c(\ell_t^{n_r}, c_t^{n_r})(1 - \tau_{\ell,t}^{n_r}) f_\ell(k_t^{n_r}, \ell_t^{n_r}) + u_\ell(\ell_t^{n_r}, c_t^{n_r}) > \gamma(1 - \overline{\tau}) f_\ell(1, R)/2 + u_\ell(\ell_t^{n_r}, c_t^{n_r}) > 0,$$

which is a contradiction. This completes the proof of Claim 1.

Since $k_1^n \to 0$, Claim 1 implies that $\lim_{n\to\infty} \ell_1^n/k_1^n = \infty$. Otherwise, there would be a subsequence $\{n_r\}$ such that $\lim_{r\to\infty} \ell_1^{n_r}/k_1^{n_r} = R < +\infty$. But this implies that $\lim_{r\to\infty} \ell_t^{n_r} = 0$, contradicting Claim 1. Since $\lim_{n\to\infty} \ell_1^n/k_1^n = \infty$, we also have that $f_k(k_1^s, \ell_1^s) = f_k(k_1^s/\ell_1^s, 1) \to \infty$.

Suppose $\limsup_{n\to\infty} k_0^n > 0$. Since $\tau_{k,0}^n \leq \overline{\tau} < 1$, either

$$\liminf_{n\to\infty}(1 - \tau_{k,0}^n)f_k(k_0^n, \ell_0^n)k_0^n > 0$$

(and thus household income at date zero is bounded away from zero) or there exists a subsequence $\{n_r\}$ such that $\ell_0^{n_r} \to 0$ (so that by (A3), $f_k(k_0^{n_r}, \ell_0^{n_r}) = f_k(k_0^{n_r}/\ell_0^{n_r}, 1) \to 0$). Either case is a contradiction. The latter implies that $\ell_0^{n_r}/k_0^{n_r} \to 0$, which contradicts Claim 1. In the former case, since $k_1^n \to 0$, $c_0^n$ is uniformly bounded away from 0. But, for $n$ large enough, the household would prefer to save more in period 0 and consume the additional income in period 1. For each additional dollar it saves, its utility decreases by at most $U_c$ in period 0. Each additional dollar invested increases the household's utility in period 1 by at least $(1 - \overline{\tau})f_k(k_1^n, \ell_1^n)\gamma$, which tends to $\infty$ as $n \to \infty$, for all $\{x_t^n\}_{t=0}^{\infty}$. Therefore $k_0^n \to 0$.

Since $k_0^n \to 0$, Claim 1 again implies that $\ell_0^n/k_0^n \to \infty$. Consumption $c_0^n$ at date zero can be expressed

$$\begin{aligned} c_0^n &= (1 - \tau_{k,0}^n)f_k(k_0^n, \ell_0^n)k_0^n + (1 - \tau_{\ell,0}^n)f_\ell(k_0^n, \ell_0^n)\ell_0^n - k_1^n \\ &\geq (1 - \tau_{k,0}^n)f_k(k_0^n, \ell_0^n)k_0^n - k_1^n \\ &> [(1 - \overline{\tau})f_k(k_0^n/\ell_0^n, 1) - 1]k_1^n \geq 0, \end{aligned}$$

where the strict inequality is a consequence of $k_0^n > \hat{k}_0^n > k_1^n$ and the last inequality is implied by $f_k(k_0^n/\ell_0^n, 1) \to \infty$. Hence, for large enough $n$, $c_0^n > 0$. As before, this leads to a contradiction: for large enough $n$, the household would prefer to save more in period 0 and consume the additional income in period 1. $\blacksquare$

PROOF OF THEOREM 1: By definition, for each $k \in [\underline{k}, \overline{k}]$ and $(m, v) \in B(W)(k)$, there exist functions $\tau_k, \tau_\ell, \ell, c, k_+, m_+, v_+ : [0, 1] \to \mathbf{R}$ such that for each $x \in [0, 1]$, the vector $\xi(x) = (\tau_k, \tau_\ell, \ell, c, k_+, m_+, v_+)(x)$ is admissible w.r.t. W at $k$ and

$$(m, v) = \int_0^1 \Psi(k, \xi(x)) \, dx.$$

Moreover, these functions can be assumed to be measurable, and so the expected value is well defined. Let $A$ be a map that for each $(k, m, v) \in \text{graph}(B(W))$ selects a function $A(k, m, v) : [0, 1] \to \mathbf{R}^7$ such that for each $x \in [0, 1]$, $A(k, m, v)(x)$ is an admissible vector w.r.t. W at $k$ and

$$(m, v) = \int_0^1 \Psi(k, A(k, m, v)(x)) \, dx.$$

22

The first coordinate of $A(k, m, v)(x)$, for example, is denoted by $A_1(k, m, v)(x)$, and as usual it represents the corresponding capital tax rate $\tau_k(x)$.

For each $(k_0, m_0, v_0) \in \text{graph}(B(W))$ we now construct a symmetric strategy profile $\hat{\sigma}(k_0, m_0, v_0)$ recursively as follows. At date $t = 0$, for each $x_0 \in [0, 1]$, let

$$\tau_0^*(x_0) = (A_1, A_2)(k_0, m_0, v_0)(x_0)$$
$$\ell_0^*(x_0) = A_3(k_0, m_0, v_0)(x_0)$$
$$y_0^*(x_0) = (1 - \tau_{k,0}^*(x_0))f_k(k_0, \ell_0^*(x_0))k_0 + (1 - \tau_{\ell,0}^*(x_0))f_\ell(k_0, \ell_0^*(x_0))\ell_0^*(x_0)$$
$$\theta_0^*(x_0) = A_5(k_0, m_0, v_0)/y_0^*(x_0).$$

Then define

$$\hat{\sigma}_G(k_0, m_0, v_0)(0)(x_0) = \tau_0^*(x_0)$$
$$\hat{\sigma}_C(k_0, m_0, v_0)(0)(x_0, \tau_0^*(x_0)) := (\ell_0^*(x_0), \theta_0^*(x_0)).$$

For $\tau_0 = (\tau_{k,0}, \tau_{\ell,0}) \neq (\tau_{k,0}^*, \tau_{\ell,0}^*)(x_0)$, let $(\ell', c', k'_+, m'_+)$ be a solution to

$$\min_{(\ell, c, k_+, m_+)} (1 - \beta)\left[u(c, \ell) + g(\tau_{k,0}f_k(k_0, \ell)k_0 + \tau_{\ell,0}f_\ell(k_0, \ell)\ell)\right] + \beta\underline{W}(k_+, m_+)$$
$$\text{s.t.} \quad (\ell, c, k_+) \in CE(k_0, \tau_0, m_+)$$

and $y' = (1 - \tau_{k,0})f_k(k_0, \ell'_0)k_0 + (1 - \tau_{\ell,0})f_\ell(k_0, \ell')\ell'$. Then, let

$$\hat{\sigma}_C(k_0, m_0, v_0)(0)(x_0, \tau_{k,0}, \tau_{\ell,0}) := (\ell', k'_+/y').$$

This specifies the period $t = 0$ strategies. For $t = 1$ and arbitrary history $h^0 = (x_0, \tau_0)$, define $(m_1, v_1)(h^0)$ as follows. If the government has not deviated (that is, if $\tau_0 = (\tau_{k,0}^*, \tau_{\ell,0}^*)(x_0)$), let

$$(m_1, v_1)(h^0) := (A_6, A_7)(k_0, m_0, v_0)(x_0).$$

If $\tau_0 \neq (\tau_{k,0}^*, \tau_{\ell,0}^*)(x_0)$, let

$$(m_1, v_1)(h^0) = (m'_+, \underline{W}(k_1, m'_+)),$$

where, as defined before, $(\ell', c', k'_+, m'_+)$ is an optimal solution to the previous optimization problem.

Note that since $A(k_0, m_0, v_0)(x_0)$ is an admissible vector w.r.t. $W$ by construction and $W \subset B(W)$ by assumption, we have that $(k_1, m_1, v_1)(h^0) \in \text{graph}(B(W))$, and thus $A((k_1, m_1, v_1)(h^0))$ is also well defined. Hence, we can specify the actions for the government and the households at $t = 1$ for each $(k_1, m_1, v_1)(h^0)$, $h_0 \in [0, 1] \times [\underline{\tau}, \overline{\tau}]^2$, in a similar fashion. Continuing this way, we can roll out the entire strategy $\hat{\sigma}(k_0, m_0, v_0)$. This recursive procedure effectively defines a family of symmetric strategy profiles, one for each $(k_0, m_0, v_0) \in \text{graph}(B(W))$.

By construction, for all $h^t = (h^{t-1}, x_t, \tau_t)$ (both on and off the path of play), $(\ell_t, c_t, k_{t+1})(h^t) \in CE(k_t(h^{t-1}), \tau_t, m_{t+1}(h^t))$. Therefore, by Lemma 3, $\hat{\sigma}(k_0, m_0, v_0)$ satisfies condition (ii) of an SSE for all $(k_0, m_0, v_0) \in \text{graph}(B(W))$.

We next show that $(m_0, v_0) = \Phi(k_0, \hat{\sigma}(k_0, m_0, v_0))$ for all $(k_0, m_0, v_0) \in \text{graph}(B(W))$. Fix $(k_0, m_0, v_0) \in \text{graph}(B(W))$, and for each $x_0 \in [0, 1]$, let

$$(\tau_{k,0}, \tau_{\ell,0}, \ell_0, c_0, k_1, m_1, v_1)(x_0) = A(k_0, m_0, v_0)(x_0) \text{ and}$$
$$G(x_0) = \tau_{k,0}(x_0)f_k(k_0, \ell_0(x_0))k_0 + \tau_{\ell,0}(x_0)f_\ell(k_0, \ell_0(x_0))\ell_0(x_0).$$

By construction

$$v_0 = \int_0^1 \{(1 - \beta)[u(\ell_0(x_0), c_0(x_0)) + g(G(x_0))] + \beta v_1(x_0)\} \, dx_0 \text{ and}$$

$$m_0 = \int_0^1 (1 - \tau_{k,0}(x_0))f_k(k_0, \ell_0(x_0))u_c(\ell_0(x_0), c_0(x_0)) \, dx_0 = \Phi_C(k_0, \hat{\sigma}(k_0, m_0, v_0)).$$

Also,

$$\Phi_G(k_0, \hat{\sigma}(k_0, m_0, v_0)) = \int_0^1 \{(1 - \beta)[u(\ell_0(x_0), c_0(x_0)) + g(G(x_0))]$$
$$+ \beta \Phi_G(k_1(x_0), \hat{\sigma}(k_1(x_0), m_1(x_0), v_1(x_0)))\} \, dx_0.$$

Hence,

$$\mid v_0 - \Phi_G(k_0, \hat{\sigma}(k_0, m_0, v_0)) \mid \leq \max_{(k_+, m_+, v_+) \in \text{graph}(W)} \beta \mid v_+ - \Phi_G(k_+, \hat{\sigma}(k_+, m_+, v_+)) \mid.$$

Since $W \subset B(W)$ and this inequality holds for all $(k_0, m_0, v_0) \in B(W)$,

$$\max \{ \mid v_0 - \Phi_G(k_0, \hat{\sigma}(k_0, m_0, v_0)) \mid \; \mid \; (k_0, m_0, v_0) \in \text{graph}(B(W)) \}$$
$$\leq \beta \max \{ \mid v_+ - \Phi_G(k_+, \hat{\sigma}(k_+, m_+, v_+)) \mid \; \mid \; (k_+, m_+, v_+) \in \text{graph}(B(W)) \}.$$

Therefore, $v_0 = \Phi_G(k_0, \hat{\sigma}(k_0, m_0, v_0))$ for all $(k_0, m_0, v_0) \in \text{graph}(B(W))$.

Finally, admissibility implies that the government constraints (condition (i) of an SSE) are satisfied. ∎

We next prove two lemmas, which combined establish Theorem 2. Fix once and for all an arbitrary usc $W$ with compact and convex values. Since $\text{graph}(W)$ is compact, there exist $\hat{m}, \underline{w}, \overline{w} \in \mathbf{R}$ such that

$$W(k) \subset [0, \hat{m}] \times [\underline{w}, \overline{w}] \quad \text{for all} \quad k \in [\underline{k}, \overline{k}].$$

LEMMA 6: *graph$(B(W))$ is bounded.*

PROOF: Pick any $k \in [\underline{k}, \overline{k}]$ and any vector $\xi = (\tau_k, \tau_\ell, \ell, c, k_+, m_+, v_+)$ admissible w.r.t. $W$ at $k$. From Lemma 2, we have that $\Psi_C(k, \xi) \in [0, \overline{m}]$.

Since $\xi$ is an admissible vector w.r.t. $W$, $(\ell, c, k_+) \in CE(k, \tau, m_+)$. Since $m_+ \geq 0$ and not working, not consuming, and not saving are feasible, household optimization implies

$$u(\ell, c) \geq u(0, 0) - \beta \, \hat{m} \, \overline{k}.$$

Hence $\Psi_G(k, \xi) \in [\underline{w}_+, \overline{w}_+]$, where

$$\underline{w}_+ := (1 - \beta)[u(0, 0) - \beta \, \hat{m} \, \overline{k} + g(0)] + \beta \, \underline{w}$$
$$\overline{w}_+ := (1 - \beta)[u(0, \overline{k}) + g(\overline{\tau k})] + \beta \, \overline{w}.$$

Therefore, $\text{graph}(B(W)) \subset [\underline{k}, \overline{k}] \times [0, \overline{m}] \times [\underline{w}_+, \overline{w}_+]$. ∎

LEMMA 7: *graph$(B(W))$ is closed.*

PROOF: Let $B^R(W)(k) = \{\Psi(k, \xi) \mid \xi$ is admissible w.r.t. $W$ at $k\}$ and thus $B(W)(k) = co(B^R(W)(k))$. Pick any sequence $\{(k^\alpha, m^\alpha, v^\alpha)\}_\alpha$ such that $(m^\alpha, v^\alpha) \in B^R(W)(k^\alpha)$ for each $\alpha$ and $(k^\alpha, m^\alpha, v^\alpha) \to (k^*, m^*, v^*)$. By definition, for each $\alpha$ there exists a vector $\xi^\alpha = (\tau_k^\alpha, \tau_\ell^\alpha, \ell^\alpha, c^\alpha, k_+^\alpha, m_+^\alpha, v_+^\alpha)$ admissible w.r.t. $W$ at $k^\alpha$, such that $(m^\alpha, v^\alpha) = \Psi(k^\alpha, \xi^\alpha)$. Since $\{\xi^\alpha\}_\alpha \subset [\underline{\tau}, \overline{\tau}]^2 \times [0, 1] \times [\underline{k}, \overline{k}]^2 \times [0, \hat{m}] \times [\underline{w}, \overline{w}]$, we can assume without loss of generality that this sequence converges to a vector $\xi^* = (\tau_k^*, \tau_\ell^*, \ell^*, c^*, k_+^*, m_+^*, v_+^*)$. By continuity, it is easy to verify that $\xi^*$ is admissible w.r.t. $W$ at $k^*$ and that $(m^*, v^*) = \Psi(k, \xi^*)$. Hence, $(k^*, m^*, v^*) \in \text{graph}(B^R(W))$ and thus $\text{graph}(B^R(W))$ is closed. Since the convex hull of a closed set is a closed set, $\text{graph}(B(W))$ is closed. ∎

The previous two lemmas establish that $\text{graph}(B(W))$ is compact. Hence $B(W)$ is usc, which establishes Theorem 2.

## REFERENCES

Abreu, D. (1986): "Extremal Equilibria of Oligopolistic Supergames," *Journal of Economic Theory*, **39**, 191 – 225.

Abreu, D. (1990): "On the Theory of Infinitely repeated Games with Discounting," *Econometrica*, **56**, 383 – 396.

Abreu, D., D. Pearce, and E. Stacchetti (1986): "Optimal Cartel Equilibria with Imperfect Monitoring," *Journal of Economic Theory*, **39**, 251–269.

Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, **58**, 1041–1063.

Atkeson, A. (1991): "International Lending with Moral Hazard and Risk of Repudiation," *Econometrica*, **59**, 1069–1090.

Benhabib, J. and A. Rustichini (1997): "Optimal Taxes Without Commitment," *Journal of Economic Theory*, **77(2)**, 231–59.

Chamley, C. (1986): "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, **54**, 607–622.

Chang, R. (1998): "Credible Monetary Policy with Long Lived Agents: Recursive Approaches," *Journal of Economic Theory*, **81(2)**, 431–67.

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

Chari, V. and P. Kehoe (1993a): "Sustainable Plans and Debt," *Journal of Economic Theory*, **61**, 230–261.

Chari, V. and P. Kehoe (1993b): "Sustainable Plans and Mutual Default," *Review of Economic Studies*, **60**, 175–195.

Judd, K. (1985): "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, **28**, 59–83.

Kydland, F. and E. Prescott (1977): "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, **87**, 473–492.

Kydland, F. and E. Prescott (1980): "Dynamic Optimal Taxation, Rational Expectations, and Optimal Control," *Journal of Economic Dynamics and Control*, **2**, 79–91.

Marcet, A. and R. Marimon (1994): "Recursive Contracts," mimeo, European University Institute.

Persson, M., T. Persson, and L.E.O. Svensson (1987): "Time Consistency of Fiscal and Monetary Policy," *Econometrica*, **55**, 1419–1431.

Rogoff, K. (1989): "Reputation, Coordination, and Monetary Policy," in *Modern Business Cycle Theory*, ed. R. Barro. Cambridge: Harvard University Press, 236 – 264.

Stokey, N. (1989): "Reputation and Time Consistency," *American Economic Review*, **79**, 134–139.

Stokey, N. (1991): "Credible Public Policy," *Journal of Economic Dynamics and Control*, **15**, 627–657.