Public Spending and Optimal Taxes Without Commitment[1]

Jess Benhabib
New York University

Aldo Rustichini
CORE

Andrés Velasco
New York University and NBER

March 11, 1998

[1]We are grateful to the C.V. Starr Center for Applied Economics at New York University for logistical support.

Abstract

We consider a representative agent, infinite-horizon economy where production requires private and public capital. The supply of public capital is financed through distortionary taxation. The optimal (second best) tax policy of a benevolent government is time inconsistent. We therefore introduce explicitly the constraint that at no point in time the revision of the original tax plan is desirable. We completely characterize the (third best) tax plan that satisfies this constraint, and estimate the difference in tax rate between the second and third best policy for a wide range of parameters. For some of these the difference between the second and third best tax rates is large, and so are the associated rates of economic growth.

1 Introduction

The question of how much government should tax and spend is one of the key policy issues of economic development. If spending is productive and taxes are distortionary, then policymakers face a well defined trade-off. The resolution of this trade-off has implications both for the level of output and the rate of economic growth.

A good starting point is the growth model developed by Barro [1]. In that model the challenge for policy is to balance the distortions to savings decisions that arise from the taxation of capital against the benefits that arise from the provision of productive public services. Barro [1] showed that for particular specifications of preferences and technology the optimal fiscal policy in that model involves a constant tax rate, which is implementable even if government cannot commit to future taxes.

In more general specifications, however, a time inconsistency problem arises, with the sequences of tax rates implemented under discretion being quite different from those (the second best policy) implemented under precommitment.[1] In this paper we therefore introduce explicitly the constraint that at no point in time the revision of the original tax plan is desirable. We completely characterize the (third best) tax plan that satisfies this constraint, and estimate the difference in tax rate between the second and third best policy for a wide range of parameters.

Consider first the government's problem under precommitment. Optimal intertemporal taxation in this context has been extensively studied since the seminal works of Chamley ([6], [7]) and Judd [11].[2] In our model, the tax problem also has clear intertemporal implications. Since future taxes affect and distort savings and consumption decisions in all of the earlier periods, under precommitment optimal tax rates will (in general) not be constant: capital taxes in the earlier periods will be less distortionary than capital taxes in the future, and therefore will present a less costly trade-off in the financing of

[1] Barro uses a Cobb-Douglas production function for output, with constant return to scale and private capital and public services as inputs. With this specification, the tax rate that maximizes output also maximizes the rate of return on capital, and therefore also maximizes the rate of growth. This simple setup has two implications: the optimal tax rate that maximizes the utility of the representative agent is constant, and there is no time inconsistency problem.

When we generalize the production function to a CES specification, the taxes that maximize output in each period, and which are the tax rates that would obtain in the discretionary equilibrium, are different than those which maximize the return on capital and the rate of growth. Furthermore, the optimal tax sequence under commitment is no longer constant (in particular, the initial-period tax rate is different from that of all subsequent periods), and the time-consistency problem emerges.

[2] In the models of Chamley and Judd the expenditures of government are modeled as an exogenous process. The tradeoff is between the intertemporal distortions of capital taxation and the distortions to labor supply decisions that come from labor taxes.

public goods.[3] This is true of in a broad class of intertemporal tax problems.[4]

Below we show that under commitment the optimal tax rates must asymptotically converge to the constant tax rate that maximizes the growth rate. This tax rate, which is independent of the capital stock because of the Barro [1] specification leading to endogenous growth, may be larger or smaller that the output-maximizing tax rate, depending on whether the elasticity of substitution between private capital and public services is above or below one. [5]

Without commitment the story is very different: the government would want to reoptimize in each period and implement the tax rates that are optimal for the initial period (in our problem, these are the tax rates and associated supply of public services that maximize single-period output). Since agents would expect this and save accordingly, the outcome would be a bad (discretionary) equilibrium with low savings.[6] The inefficiency of such discretionary equilibria has led to a literature on "reputational equilibria" along the lines of trigger strategy equilibria in game theory. In models with reputational equilibria, the government must weigh the one-time benefits of deviating from the announced optimal policies against a loss of reputation that leads to a switch to policies and actions associated with the discretionary (bad) equilibrium. Many authors have shown, in a variety of contexts, that such reputational mechanisms can often sustain policies that are optimal under commitment.[7]

But it is also possible –especially in a model with capital accumulation–

[3]In fact, since in the initial period capital is in fixed supply the optimal tax strategy will be to tax capital at a rate high enough that will force the public to borrow from the government a quantitity sufficient to generate interest income that will pay for all of its optimally desired expenditures. Since such a scheme with a negative net value of government bonds is highly unrealistic, it is standard in the optimal tax literature to impose bounds on the (negative) government bonds, or to limit the maximal tax rate in the first period. In the same spirit, we adopt in this paper a simplifying assumption to rule out government lending or borrowing, and we will require a balanced budget. The tradeoffs inherent in the problem of optimal capital taxation will remain, and be little affected under this assumption. The case with bonds in a model of the type considered by Chamley and Judd, but without commitment, is presented in Benhabib and Rustichini [3].

[4]In fact, Chamley and Judd obtain the remarkable result that if in the limit (as time tends to infinity), the economy converges to a stationary state, optimal capital tax rates under commitment must approach zero.

[5]In Benhabib and Velasco [4], a similar question is studied in the context of an open economy with international capital mobility. There it is assumed that capital taken abroad can avoid domestic taxes. Since world interest rates are fixed from the perspective of a small country, the analysis is considerably simplified. In contrast to the results of this paper, the optimal taxes under commitment, as well as those without commitment, turn out to be constant after the initial period.

[6]Results of these type are of course examples of a broader "time-consistency" problem discussed by Kydland and Prescott [12].

[7]See for example Stokey [16], Chari and Kehoe [9] or in a pure monetary context Barro and Gordon [2]. Lucas and Stokey [13] go a different route, studying how the maturity of government debt might render the commitment tax sequence time-consistent.

that the one-time advantage of deviating from announced policies is so large that commitment policies are not sustainable, even if the consequence is the loss from deviating is a permanent loss of governmental reputation. Such situations lead quite naturally to the question of what the are the best sustainable or time-consistent tax policies. In this paper we fully characterize the best sustainable taxes and levels of public services in circumstances where a commitment by the government to future policies is not possible.

To characterize fully this best sustainable tax sequence we proceed in two steps. First, we must identify and characterize the worst possible equilibrium, which is the one to which the economy will revert if a deviation from the announced path takes place. It turns out that, in the extended Barro model, the worst (perfect) equilibrium is one in which, with a probability $1 - \varepsilon$ that is arbitrarily close to one, the government myopically implements the tax rate that maximizes output; since this tax rate is less than one, agents provide positive savings, and the economy grows over time. Notice that, unlike the Chari-Kehoe [8] model, this worst equilibrium is not the autarkic one in which the government attempts to tax all capital away and agents therefore save nothing. We provide below some intuition for this crucial difference between the two models.

Second, we compute the *third best* tax sequence that can be sustained by the threat of reversion to that worst equilibrium. Under some parametrizations the optimal tax rates associated with the commitment outcome cannot be sustained because the value of deviation at some future period exceeds the value of continuation under the announced policy. We show that, in that case, the best sustainable asymptotic tax rates are constant and must lie between the output-maximizing and growth-maximizing tax rates. In particular, we show that in some very reasonably calibrated examples the optimal capital taxes under commitment can be substantially smaller than the best sustainable capital taxes.

The model presented above can be viewed within the broader context of the problem of characterizing best sustainable equilibria in dynamic games. Restricting attention to the best sustainable equilibria rather than all sustainable equilibria allows us to formulate the problem as an optimization problem subject to period-by-period incentive compatibility constraints that require the value of continuation to be larger than the value of deviation. Problems of this type have been considered by Marcet and Marimon [14] and by Benhabib and Rustichini [3]. In the optimal taxation problem considered in this paper, one further simplification arises from the Stackelberg nature of the game: the government moves first. On the other hand, additional difficulties arise from the fact that current saving decisions are complicated functions of all future tax rates.

The model also has implications for economic growth. The second best (commitment) solution would involve setting the tax rate so as to maximize the growth rate of the economy in the long run. Such a policy, however, is time inconsistent. The third best (reputational) policy involves a tax rate that is a compromise between the twin objectives of maximizing the rate of growth and maximizing the current period's level of output –and which therefore generally does not maximize the rate of growth. In short, time inconsistency leads to lower economic growth.

The next section describes the model and the equilibrium. Section 3 discusses the optimal taxes under commitment. Later sections identify and characterize the worst perfect equilibrium. The section after that sets up the problem with incentive constraints when commitment by the government to future tax rate is not possible. Section 6.2 provides some particular examples and provides a characterization of the optimal taxes without commitment. Section 7 provides a family of calibrated examples to illustrate numerically the differences between optimal taxes with and without commitment, while Section 8 concludes.

2 An Economy with Private and Public services

We first describe our simple economy. There is one representative agent, who has an infinite life, and one private good, which is used both in production and in consumption. There are also public services, which are only used in production, and for which the government does not charge private agents. The government is benevolent, and can freely choose tax rates for the purpose of maximizing the utility of the representative agent. Taxes are used to finance the provision of the public capital good.

The technology

At the end of each period a certain amount of private capital, k_t, is available. Out of this a total amount $\tau_t k_t$ is then taxed away from the agent, and the rest is available for production. The output of the good is determined by a CES production function where both private capital and public services enter:

$$y_t = A\left(a(1-\tau_t)^{-\rho}k_t^{-\rho} + (1-a)g_t^{-\rho}\right)^{-\frac{1}{\rho}} \tag{2.1}$$

where g_t is the flow of public services, $\rho \in [-1, +\infty]$, $A \geq 0$. Note that the tax rate τ appears in the production function because $(1-\tau)\,k$ is the amount of private capital left untaxed and therefore available for production. The tax rate τ is usually assumed to range in the interval $[0,1]$, but in some special case we may want to consider the restriction $\tau \in [\tau_L, \tau_B]$.

The provision of public services takes a very simple form. The total amount of taxes is converted one-to-one into public good, so that

$$g_t = \tau_t k_t \tag{2.2}$$

Using 2.2 in 2.1 we have

$$y_t = Ak_t \left(a(1-\tau_t)^{-\rho} + (1-a)\tau_t^{-\rho}\right)^{-\frac{1}{\rho}} \equiv Ak_t\phi(\tau_t) \tag{2.3}$$

Hence, the amount of private capital k_t affects total output in two ways: directly, and indirectly through the effect on total taxes and hence on the total amount of public services.[8]

In evaluating the return on private capital the representative agent will take the amount g_t as given, and will ignore the indirect effect. In equilibrium, however, the condition $g_t = \tau_t k_t$ can be substituted into the partial derivative of output with respect to k to obtain the marginal return that the agent is facing. If we do this we have (suppressing time subscripts for τ):

$$\frac{\partial y_t}{\partial k_t} \equiv AR(\tau) = Aa\phi(\tau)^{1+\rho}\left(1-\tau\right)^{-\rho} \tag{2.4}$$

Where clearly $R\left(\tau\right) > 0$ for every τ.[9] Note that

$$\begin{aligned} R(\tau) &= a\phi(\tau)^{1+\rho}\left(1-\tau\right)^{-\rho} \\ &\leq a\phi(\tau)^{1+\rho}\left(1-\tau\right)^{-\rho} + (1-a)\phi(\tau)^{1+\rho}\tau^{-\rho} \\ &= \frac{\partial\phi}{\partial(1-\tau)}(1-\tau) + \frac{\partial\phi}{\partial\tau}\tau = \phi(\tau) \end{aligned} \tag{2.5}$$

by the homogeneity of degree one of $\phi\left(\tau\right)$ in $(1-\tau)$ and τ. In particular, we have:

$$R(\tau) \leq \phi(\tau), \tag{2.6}$$

an inequality that we shall use frequently later. Note that if $\rho = 0$ we have Cobb-Douglas production, so that $R/\phi = a$.

[8] Note that, because we are operating in a representative gent world, we have made no distiction between aggregate and per capita provision of public goods. In a world with many agents, on the other hand, that distinction would matter, ans would the distinction betwen rival and non-rival public services. The essentials of the analysis that follows would survive any of these extensions.

[9] If one wanted to ensure that clearly $R\left(\tau\right) > 1$ for every τ, it is possible to add a term Bk_t (where $B > 0$ is a parameter) to the production function to capture some pre-tax return on private capital. This would not affect any of the results that follow. Alternatively, one could confine the tax rate τ to a closed interval in $(0,1)$ and make A large enough: this would also ensure the result $R\left(\tau\right) > 1$ for all allowable tax rates.

Preferences and budget constraints

The representative agent owns the private capital stock. In each period he decides how much to consume out of the return on the private capital, minus the taxes plus a government transfer. The individual budget constraint has the form:

$$k_{t+1} = R(\tau_t)k_t + M_t - c_t = \phi(\tau_t)k_t - c_t \tag{2.7}$$

where M_t is the government transfer of the residual output after payments to capital.

The agent's utility from a consumption stream $\{c_t\}_{t\geq 0}$ is given by:

$$\sum_{t=0}^{\infty}(\frac{\sigma}{\sigma-1})c_t^{\frac{\sigma-1}{\sigma}}\beta^t \tag{2.8}$$

The competitive equilibrium for an given tax sequence

We now fix a sequence of tax rates $\tau = (\tau_0, \tau_1, \ldots)$ and compute the equilibrium of the economy for this arbitrary sequence. Details of the analysis are in the Appendix. The essential result is that for a given initial capital stock k_0 and a sequence of tax rates τ, lifetime utility for the agent can be written as a function $V(k_0, \tau)$, so that utility depends only on initial capital and the sequence of tax rates. The function is found to have the simple form

$$V(k_0, \tau) = (\frac{\sigma}{\sigma-1})\,(k_0)^{\frac{\sigma-1}{\sigma}}\,(\phi(\tau_0))^{\frac{\sigma-1}{\sigma}}\,h\,(\tau_1, \tau_2, \ldots) \tag{2.9}$$

Note that the initial capital stock k_0 and the sequence of tax rates $(\tau_0, \tau_1, \ldots)$ factor in the expression defining the value. In fact, it will be useful to isolate the term which involves only the initial and future expected tax rates:

$$H(\tau_0, \tau_1, \ldots) \equiv \phi(\tau_0)^{(\frac{\sigma-1}{\sigma})}h(\tau_1, \ldots) \tag{2.10}$$

so that utility takes the even simpler form

$$V(k_0, \tau) = (\frac{\sigma}{\sigma-1})\,(k_0)^{\frac{\sigma-1}{\sigma}}\,H(\tau_0, \tau_1, \ldots) \tag{2.11}$$

Notice that when the terms of the sequence τ are constant over time then H has a simple form

$$H(\tau, \tau, \ldots) = (\phi\,(\tau) - (\beta R(\tau))^{\sigma})^{\frac{\sigma-1}{\sigma}}\left(1 - \beta^{\sigma}R(\tau)^{\sigma-1}\right)^{-1} \tag{2.12}$$

For convenience, we assign a special symbol to the following function defined on $[0, 1]$:

$$H^*(\tau) \equiv H(\tau, \tau, \ldots). \tag{2.13}$$

This function will prove useful in the analysis that follows.

3 The optimal tax with commitment

In the previous section we have reduced the value to the agent, at the competitive equilibrium for a given sequence of tax rates, to a closed form expression in terms of the tax rates. If the government can commit to a sequence of tax rates in the first period, without possibility of revising the decision later, then the optimal sequence of tax rates is easy to determine. Existence of the optimal tax in our model is proved in the Appendix.

The trade-off between higher and lower tax rates should be clear. Since agents in the economy have no incentive to contribute to the accumulation of the public good, taxes are essential to its provision. The taxes that allow for the provision of public services, which is a complement in production, may well reduce the incentives for to private accumulation.

The details of the analysis of the optimal tax policy with commitment are in the Appendix. The main conclusion we derive there is that a necessary condition for the optimal tax sequence is that

$$\lim_{t \to \infty} R'(\tau_t) = 0 \tag{3.14}$$

That is to say, in the limit the optimal tax rate is that which maximizes the agent's perceived marginal return on investment, and therefore also maximize growth. Notice that $\widetilde{\tau}$ is henceforth defined as the tax rate that solves $R'(\widetilde{\tau}) = 0$. Under the assumed CES technology, the tax rate given by 3.14 is different from the tax rate given by $\phi'(\bar{\tau}) = 0$ which maximizes output within the period, but induces distortions to the accumulation process.

There is a clear similarity between condition 3.14 and the Chamley-Judd (see [6], [7], [11]) result that in the limit the tax rate on capital income goes to zero in optimal taxation with commitment: in both cases the tax rates in question maximize the growth rate of the economy.

In the case of Cobb-Douglas technology ($\rho = 0$) studied by Barro [1], we have an especially simple case. Since $R(\tau)/\phi(\tau) = a$, so that $\widetilde{\tau} = \overline{\tau}$. In that case, it is clear by inspection that $R'(\tau_t) = 0$ for all $t > 0$. We also have that $\phi'(\tau_0) = 0$, from maximizing $V(k_0, \tau)$. So τ_t is constant starting at time zero, and equal to the rate $\widetilde{\tau} = \overline{\tau}$. In words, the tax rate is constant and equal to the rate that simultaneously maximizes output levels and output rates of growth.

4 The Model as a Game

We now dispose of the assumption that the government can precommit, and study the determination of tax rates as the outcome of a game between the government and the public. To do so we must be a great deal more specific about the strategies open to both the government and the public, and about the equilibrium concept.[10]

4.1 Histories and Strategies

The set of histories is the union of two distinct set of histories, one relevant for the government and the other relevant for the agents in the economy. So we let:

$$H \equiv H_G \cup H_P. \tag{4.15}$$

The histories in H_G have the form $(k_0, \tau_0, \ldots, k_{t-1}, \tau_{t-1})$; those in H_P have the form $(k_0, \tau_0, \ldots, \tau_t)$. Note that we do not report in the histories the savings of each individual, but only the aggregate savings. We call these "public histories".

A strategy for the government is a sequence $\gamma \equiv (\gamma_0, \ldots, \gamma_t, \ldots)$, where each γ_t is a function from H_G to the unit interval, describing the tax rate on capital for the period t as a function of the public history.

A strategy for the each agent in the public is a sequence $\pi^i \equiv (\pi_0^i, \ldots, \pi_t^i, \ldots)$, where each π_t^i maps the history of past tax rates and savings into future saving.

In our model the description of the strategies is simple. In fact, in this model agents are negligible, so the variables which are relevant for their decisions – namely the future tax rates and the future values of the aggregate capital– are independent from the actions of each single agent. Each agent cannot affect the value of the aggregate capital, since she is negligible, and she does not therefore have any influence on the tax rates, which depend on the aggregate savings. This implies that for a given strategy of the government the strategy of the public can be derived by solving the maximization problem of each agent. This fact is stated in the following lemma.

Lemma 1 *For each strategy of the government γ there is an induced strategy π of the public.*

Proof. Take for any t any history in H_P , and determine the pair of sequences $(k_s^*, \tau_{s+1}^*)_{s \geq t}$, such that

i. the sequence of taxes is the one induced by γ and the sequence $(k_s^*)_{s \geq t}$;

[10] In setting up the model as a game we follow Chari and Kehoe [8], adjusting the structure to our situation. The main difference is that we model explicitly the strategy and the choices of the public (as they do in the section V of the paper, "Anonymous games", but not ins the main sections of the paper). The reason for this change is that we want to analyze the robustness of the equilibria in the game. To do this we need to introduce the strategies of the public explicitly. An alternative approach could be of course to consider small mistakes in the allocation rule used by the public. We chose this alternative because it is easier to use directly refinement concepts from game theory.

ii. the sequence of capital stocks is determined in each period by the maximization of the utility of each player, given the sequence $(k_s^*, \tau_{s+1}^*)_{s\geq t}$ of future aggregate capital and tax rates.

4.2 Definition of Equilibrium

In our model, a *sustainable equilibrium* is defined as follows:

i. a history at t is a sequence $(\tau_0, \ldots, \tau_{t-1})$ of tax rates;

ii. a policy plan for the government is a function γ from histories into tax rates;

iii. an allocation rule is a function from histories to savings that solves the consumer's problem.

5 Perfect equilibria

We now characterize the worst perfect equilibrium in the Barro [1] model, which is the one will use as a threat in constructing the best sustainable tax sequence.

In this model there are several strategy profiles that can give rise to "bad" equilibria. By analogy with the Chari-Kehoe [8] model, one might expect the following strategies to be an equilibrium: the government sets $\tau_t = 1$ in every period, irrespective of the history. Each agent i sets the saving $k_t^i = 0$ in every period, also irrespective of the history. As the Appendix shows, this strategy profile gives rise to an equilibrium (termed Bad Equilibrium 1) that is both Nash and a sustainable equilibrium (the latter in the sense of Chari-Kehoe [8]), but it is not subgame perfect.

An alternative candidate strategy profile is the following: government chooses a tax rate equal to one in any period in which the aggregate savings are equal to zero; it chooses $\bar{\tau}$ in every period in which aggregate savings are positive; agents choose savings equal to zero in every period. As the Appendix also shows, these strategies give rise to an equilibrium (termed Bad Equilibrium 2) that is subgame perfect. However, it is not robust to small mistakes by players –that is, it is not a perfect equilibrium.

There is a third strategy profile, nevertheless, that does give rise to a perfect equilibrium. Informally, this is a situation where the government always implements the myopic output-maximizing tax rate ($\gamma_t(h_t) \equiv \bar{\tau}$ for every t and every history h_t) and the public follows the associated induced strategy, is a perfect equilibrium.

The technical definition of perfect equilibria here is complex, because there is a continuum of agents, and the action set of each has uncountable cardinality. To adapt the definition of perfect equilibria to this situation, we define a totally mixed strategy as a strategy which, with probability $\epsilon > 0$, chooses a value that is in turn drawn from a uniform distribution over the strategy space of the player (the unit interval for the government, the set of feasible savings for the agents).

Strategies are:

- the government sets the tax rate in each period to $\bar{\tau}$ with probability $1-\epsilon$. (The value of the tax rate is precisely $\bar{\tau}$, and is not affected by the ϵ);
- the public expects the tax rate $\bar{\tau}$ in each period with probability $1-\epsilon$, and with probability ϵ they expect the tax rate to be a random draw from a uniform distribution over the unit interval ; they maximize accordingly.

The ϵ *does* matter for the value of this program; so rather than a function $V(k)$ we have a function $V(k, \epsilon)$. However:

$$\lim_{\epsilon \to 0} V(k, \epsilon) = V(k) \tag{5.16}$$

for every k, even uniformly.

After any deviation, the government sets tax rate $\bar{\tau}$, also with probability $1-\epsilon$; the public saves optimally for a given sequence $\bar{\tau}$ of tax rates.

Theorem 1 *The "third best" strategies described above are a perfect equilibrium. The strategy of the government*

$$\gamma_t(h_t) \equiv \bar{\tau}$$

and the induced strategy for the public give rise to the worst Perfect Equilibrium;

Proof. For any history $h_t \in H_G$, the worst possible equilibrium depends only on the aggregate value of capital, and for any history $h_t \in H_P$ the worst possible equilibrium depends only on the aggregate capital and on the tax rate.

Notice also that the value at the worst equilibrium is increasing in the value of the aggregate capital stock. Now consider the problem of determining the tax rate, for any $\epsilon > 0$. Considering that after the period t the public and government will follow the continuation of the worst equilibrium, the government is maximizing the value of the utility of the public by choosing $\bar{\tau}$, which will give the largest value of the capital stock in the next period. Hence with probability $1-\epsilon$ the government will choose a tax rate equal to $\bar{\tau}$. ■

There seems to be a considerable difference between our setup and the one of Chari-Kehoe, in view of what we have seen in this section. The single most important difference is that in Chari-Kehoe the autarchy equilibrium in which the tax rate on capital is one and private savings are zero is not only subgame perfect, but it can also easily be shown to be robust in the sense of being perfect. By contrast, in our setup an equilibrium in which the tax rate is one is not perfect –and the worst perfect equilibrium is one in which the tax rate is $\bar{\tau}$.

6 Characterizing Third Best Tax Rates

We now explicitly characterize the third best tax sequence that is supported by reputational considerations. By the definition of a perfect equilibrium, we let ε, the probability of mistakes by the government, go to zero. That means that, after deviation from the announced tax sequence, the economy reverts to a situation in which the government implements the output maximizing tax rate $\bar{\tau}$ with probability one, and the public saves accordingly.

6.1 The value of deviation

We begin by characterizing the value of deviating from the third best tax sequence. We can characterize this value before having characterized that tax sequence because, in this model, the value of deviation is independent of the path from which the government is deviating. That is not true in other models.

After the government deviates from a previously announced plan of tax rates, the sequence of events is the following. The public will believe that from that period on the government will maximize per period output –that is, will choose in each period the constant tax rate that maximizes the function $\phi()$. Appendix 8 shows that, if the public holds such expectations, the government's best response is to fulfill them starting the period after the deviation.

Anticipating this punishment, a government contemplating the possibility of deviating will choose the optimal deviation, which will in fact consist of maximizing output the period of deviation as well. The competitive equilibrium associated with this sequence of events is easy to characterize.

If agents expect all future values of τ to be set at $\bar{\tau}$, then it can be shown (details of this computation are in Appendix 8) that the closed form solution for the value of deviation is

$$V^D(k_t) = \left(\frac{\sigma}{\sigma - 1}\right)(k_t)^{\frac{\sigma-1}{\sigma}} H^*(\bar{\tau}) \tag{6.17}$$

where

$$H^*(\bar{\tau}) \equiv \left(\phi(\bar{\tau}) - (\beta R(\bar{\tau}))^{\sigma}\right)^{\frac{\sigma-1}{\sigma}} \left(1 - \beta^{\sigma} R(\bar{\tau})^{\sigma-1}\right)^{-1} \tag{6.18}$$

Let us point out immediately the features of this expression that we shall need in what follows, the value of deviating is independent of the current announced tax rate; results from the fact that when deviation begins, the government uses the $\bar{\tau}$ rate immediately. Second, the term $H^*(\bar{\tau})$ is a constant, and the dependence on the capital is of a very specific form –that is, it has the same functional form as utility. This will make a nice cancellation possible in the next step.

Take in fact, for a given tax plan τ, and a the capital stock k_t at period t the value to the representative agent of continuing with the plan, $V^C(k_t, \tau)$

$$V^C(k_t, \tau) = \left(\frac{\sigma}{\sigma - 1}\right) (\phi(\tau_t)k_t)^{\frac{\sigma-1}{\sigma}} h(\tau_{t+1}, \tau_{t+2}, ...) \tag{6.19}$$

A tax plan is credible if and only if it does not give an incentive to the government to a revision in any period; that is, if and only if in each period after the first the continuation value $V^C(k_t, \tau)$ is larger than the deviation value $V^D(k_t)$.

But given the specific form of these two functions, the difference

$$V^C(k_t, \tau) - V^D(k_t) = \left(\frac{\sigma}{\sigma - 1}\right) k_t^{\frac{\sigma-1}{\sigma}} \left((\phi(\tau_t))^{\frac{\sigma-1}{\sigma}} h(\tau_{t+1}, \tau_{t+2}, ...) - H^*(\bar{\tau})\right) \tag{6.20}$$

is non negative if and only if

$$(\phi(\tau_t))^{\frac{\sigma-1}{\sigma}} h(\tau_{t+1}, \tau_{t+2}, ...) - H^*(\bar{\tau}) \geq 0, \text{ for all } t \geq 1 \tag{6.21}$$

The remarkable feature of 6.21 is that is does not involve the value of capital, but only of the future string of tax rates.

Before we proceed to analyze the general case, let us consider what 6.21 implies in the case of a Cobb-Douglas technology. When $\rho = 0$ then (as we have discussed above) in the optimal tax sequence for the commitment case τ_t is equal to a constant for every $t \geq 0$. So the infinitely many constraints of 6.21 reduce to a single constraint. In the general case, however, we have to take into account the entire (infinite) set.

6.2 The third best problem

If the government, at the moment of choosing a tax plan, has no commitment power over its own choices, then it has to confine the choice to the set of tax plans that in each period satisfy the constraint 6.21, which we call the *Incentive Compatibility Constraint*. Let us specify formally the problem that the government must solve. We have seen in the previous section that the value of the capital stock can be factored out in all of the constraints. So the problem

$$\max_{\tau=(\tau_0, \tau_1, ...)} V(k_0, \tau) \tag{6.22}$$

$$\text{subject to } V^C(k_t, \tau) \geq V^D(k_t) \text{ for all } t \geq 1 \tag{6.23}$$

is equivalent to the following simple problem, which is independent of the initial capital stock:

$$\max_{\tau=(\tau_0, \tau_1, ...)} \left(\frac{\sigma}{\sigma - 1}\right) H(\tau_0, \tau_1, \ldots) \tag{6.24}$$

$$\text{subject to } \left(\frac{\sigma}{\sigma-1}\right) H(\tau_t, \tau_{t+1}, \ldots) \geq \left(\frac{\sigma}{\sigma-1}\right) H^*(\tau_p) \text{ for all } t \geq 1. \quad (6.25)$$

This problem has a Lagrangean:

$$L(\tau, \lambda) \equiv H(\tau_0, \tau_1, \ldots) + \sum_{t=1}^{\infty} \lambda_t \left[H(\tau_t, \tau_{t+1}, \ldots) - H^*(\tau_p) \right]. \quad (6.26)$$

where the λ's are the Lagrange multipliers.

The necessary condition for the (third best) interior solutions is

$$\frac{\partial L}{\partial \tau_m} = 0 \text{ for every } m \geq 0. \quad (6.27)$$

We now use this condition to characterize the limit tax rates in the constrained optimal (third best) problem.

Incentive compatible limit tax rates

In the commitment case, the limit tax rate is defined (as we have seen previously) by the equation:

$$R'(\tilde{\tau}) = 0. \quad (6.28)$$

The tax plan with such a limit tax rate may or may not be incentive compatible. In Section 4 above we saw that for a tax sequence (equivalently, tax plan) to be incentive compatible, the value of continuation along that sequence must be no smaller than the value of deviation. How does such a criterion perform in the limit case we are considering?

In fact as $t \to \infty$, the difference between the continuation value from the tax plan and the value from deviation tends to a positive multiple of

$$H^*(\tilde{\tau}) - H^*(\bar{\tau}) \quad (6.29)$$

If this difference is strictly negative then the second best (commitment) tax plan is not incentive compatible, because the incentive constraint will be violated at some date in the future. In a later section we present several examples and numerical estimates of cases in which indeed the commitment plan is not incentive compatible.

The third best tax rate

When the second best optimal tax rate is not incentive compatible in the limit, we want to characterize the limit tax rates for the optimal constrained solution. They can be determined by a very simple procedure.

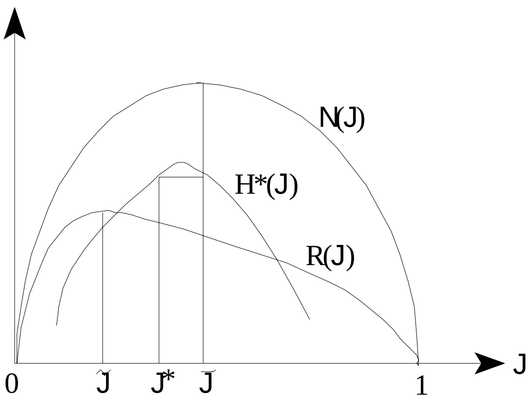

Figure 1: $\rho < 0$

First, two cases are possible: $\tilde{\tau} < \bar{\tau}$ or $\tilde{\tau} > \bar{\tau}$. We find that $\tilde{\tau} < \bar{\tau}$ if and only if $\rho < 0$. One has in fact that

$$\phi'(\tau) = \phi(\tau)^{1+\rho}[(1-a)\tau^{-(1+\rho)} - a(1-\tau)^{-(1+\rho)}]; \bar{\tau} = [1 + \left(\frac{a}{1-a}\right)^{\frac{1}{1-\rho}}]^{-1} \quad (6.30)$$

and

$$R'(\bar{\tau}) = a\rho\phi(\bar{\tau})^{1+\rho}(1-\bar{\tau})^{-(1+\rho)} \quad (6.31)$$

so if $\rho < 0$ then $R'(\bar{\tau}) < 0$, and $\tilde{\tau} < \bar{\tau}$; the converse is true if $\rho > 0$.

We shall see that in the case in which the incentive constraints are binding, the constrained optimal tax rate τ^* is a compromise between the unconstrained tax rate and the output-maximizing rate. So the constrained tax rates are higher than the unconstrained if $\rho < 0$, and lower in the converse case. These two cases and the corresponding relations between $\tilde{\tau}, \bar{\tau}$ and τ^* are illustrated in Figures 1 and 2. Note that the function $H^*(\tau)$ is not drawn to scale relative to $R(\tau)$ and $\phi(\tau)$: what matters and follows from the analysis is its relative position, so that τ^* is between $\bar{\tau}$ and $\tilde{\tau}$.

These two cases will turn out to be symmetric, although of course with very different implications from the point of view of economic analysis. For the sake of brevity in what follows we concentrate on the first case, where $\rho < 0$.[11]

Obviously, the interesting case is now $H^*(\tilde{\tau}) < H^*(\bar{\tau})$; *i.e.* the return after tax on capital is too high in the limit, and deviation to the output-maximizing tax becomes a dominant choice. This inequality implies that H^* is increasing in the interval $[\tilde{\tau}, \bar{\tau}]$.

[11] There is a converse where all arguments that follow hold, but with reversed inequalities.

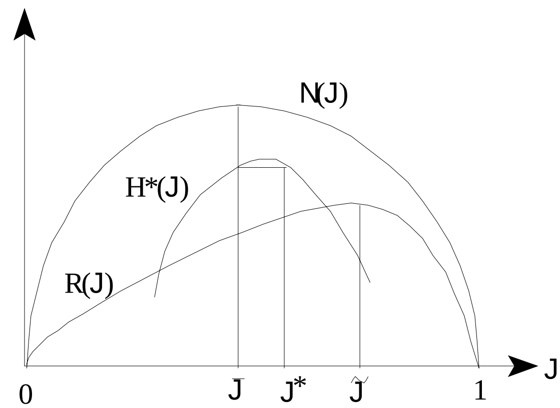

Figure 2: $\rho > 0$

Now define τ^* formally to be

$$\min\{\tau \in [\tilde{\tau}, \bar{\tau}] : H^*(\tau) = H^*(\bar{\tau})\} \tag{6.32}$$

That is to say, $\tau*$ is the lowest tax rate in between $\tilde{\tau}$ and $\bar{\tau}$ such that the value of continuation and deviation are equal.

Consider the interesting case where $\tau^* < \bar{\tau}$. Then we have immediately that

$$\frac{dH^*}{d\tau}(\tau^*) \geq 0, \text{ and } \frac{dR}{d\tau}(\tau^*) \leq 0 \leq \frac{d\phi}{d\tau}(\tau^*). \tag{6.33}$$

We now claim that at τ^* the necessary condition for optimality (in the limit) are satisfied: that is, that Lagrange multipliers exist for which τ^* satisfies the corresponding necessary conditions. The discussion of these details is technical, and can be found in the Appendix.

In the converse case of $\rho > 0$, the constrained tax rate is smaller than the limit value $\tilde{\tau}$ of the commitment solution. This difference should not hide, however, the fundamental similarity in the adjustment mechanism: in both cases the limit continuation value is too small, and the adjustment in the limit value of the tax rate makes it large enough to prevent deviation. In both cases this is achieved by increasing the limit value of output per period, at the expense of the limit value of the growth rate.

Note that, in the limit, implementing τ^* yields the same value as implementing $\overline{\tau}$ from that point on. That is to say, thereafter the value of the third-best is equal to the value of deviation associated with implementing the constant $\overline{\tau}$ forever. One might wonder, then, what is gained from implementing this third-best tax rate. Note also from Figures 1 and 2 that there is another constant τ that lies between $\overline{\tau}$ and τ^* –call it τ^H– that maximizes $H^*(\tau)$ and that would yield even higher utility from that point on. One might also wonder why it is not preferable to implement this intermediate tax rate in the limit.

The answer to both questions is the same: from the perspective of the initial planning period, it would not be optimal to implement these tax rates unless we constrain the optimization to constant sequences. Implementing $\overline{\tau}$ from the start would lower expected rates of return on investment and would deter capital accumulation. Implementing τ^H would yield a value such that the incentive compatibility constraint is not binding in the limit, so that this tax rate is implementable. We know, however, that the unconstrained optimum sequence of taxes converges to $\widetilde{\tau}$, which maximizes $R(\tau)$. Therefore, τ^H cannot be the optimal asymptotic tax rate from the perspective of the initial period. More formally, we show in the Appendix that if tax rates converge to a constant τ, the Lagrange multiplier associated with the incentive constraint must be bounded away from zero in the limit (see the condition 0.68). Since implementing τ^H in the limit yields $H\left(\tau^H\right) > H(\overline{\tau})$, the incentive constraint is not binding and the associated Lagrange multiplier must be zero, in contradiction to the condition 0.68 given in the Appendix.

It may be useful to contrast the resolution of the incentive compatibility problem in the present model and in the analogue of the Chamley-Judd model of optimal taxation with two factors of production (see Benhabib and Rustichini (1995)). The similarity is in the fact that in both models we have cases where the constrained taxes on capital are different than the unconstrained ones. The difference is in way the incentive compatibility problem is resolved. In the present model the amount of capital (on the balanced growth path) is not affected by the rate of return; similarly the value of defection is not dependent on the capital stock. All the adjustment has to take place by variations in the tax rates. In the two-factor model, the driving force of the adjustment is the following: *adjust the rate of return on capital in order to adjust the long run supply of capital, which in turn deters deviation.*

7 Second and third best: numerical values

In this section we compute a family of calibrated examples to illustrate the divergence between the limiting optimal tax rates under commitment and the optimal tax rates that are time consistent. We also show how these tax rates change with the parameter a, which measures he relative importance of private capital in production, and how they vary with the parameter ρ, which measures the elasticity of substitution between private capital and public services. In all cases we set $A = 30$; we also choose $\sigma = 0.5$ and $\beta = 0.95$, very much standard values.

The following table illustrates the numerical results discussed above for various values of a and ρ. The first number in each cell represents the second-best tax that could be enforced under commitment, while the second number is the sustainable (third-best) taxes that are incentive constrained.

$\rho \backslash a$	0.3	0.5	0.8
-0.5	.303, .408	.146, .394	.015, .055
-0.1	.695, .699	.495, .498	.196, .197
0.1	.705, .701	.505, .502	.204, .203
0.5	.804, .706	.642, .539	.369, .295
2.0	.729, .646	.654, .560	.522, .421
10.0	.592, 566	.572, .545	.539, .508

First we note, as expected, that tax rates are higher when the coefficient measuring the relative importance of the public good, $(1-a)$, is higher. As we vary ρ, taxes increase as we approach and cross the Cobb-Douglas value $\rho = 0$. As we further increase ρ and approach a fixed coefficient Leontief technology, both the private and public good become increasingly essential to production and the tax rates decline again towards the 50% rate. The difference between the second best and third best tax rates declines. In the limit, as we approach the Leontief technology, the second best tax sequence becomes sustainable. For example, when $a = 0.5$ and $\rho = 100$, the incentive constraints are satisfied for the sequence of second best tax rates. Note also that at values of ρ close to 0 (the Cobb-Douglas value) the second and third tax rates are very close. As noted before, this is because for a Cobb-Douglas technology the tax rate that maximizes the rate of return R and the tax rate that maximizes output are the same: there is no time inconsistency. Any slight deviation from Cobb-Douglas does create a time inconsistency problem, however. For all parametrizations presented here, the sequence of second best (commitment) tax rates is not sustainable and differs from the best sustainable sequence.

The difference in the asymptotic second and third best tax rates are often quite significant for the parametrizations presented above. In the case of a low elasticity of substitution, for example ($\rho = 2$), the difference in the two tax rates is 10% when $a = 0.8$. The difference is also 10% when the elasticity of substitution is large ($\rho = -0.5$) and $a = 0.3$. As expected, the differences are smaller when the technology is close to Cobb-Douglas.

The table below illustrates the resulting ratios of government expenditures to total output. As in the table above, the first number in each cell corresponds to the outcome under commitment, while the second number corresponds to the outcome under the sustainable (third best) tax rate. These ratios of government expenditures to total output are sensitive to the choice of the constant term A in the production function, since the tax rate τ primarily depends on the share parameter a and the elasticity parameter ρ.

$\rho\backslash a$	0.3	0.5	0.8
-0.5	.024987, .029590	.011409, .026569	.000746, .002696
-0.1	.042377, .042604	.033001, .033200	.010611, .010666
0.1	.043641, .043375	.033669, .0.33467	.011389, .013340
0.5	.056994, .045232	.045596, .036098	.021966, .017160
2.0	.056479, .043449	.050403, .038151	.035809, .026309
10.0	.043116, .039130	.041788, .037825	.038308, .034224

The ratios average about 3% for our parametrizations. If we interpret the numbers as the share of public investment (infrastructure, etc.) in output, they are quite realistic.

8 Conclusions

We consider an economy where a benevolent government faces a trade-off between supplying productive public services and taxing private capital and thereby distorting savings and investment decisions. When the optimal path of taxes under precomittment (the second best path) turns out to be time inconsistent, reputation may enable government to sustain a third-best tax path. We fully characterize this third best policy, which asymptotically turns out to consist of a constant tax rate that lies between the precomittment (second best) and the discretionary tax rates. Simulations for plausible parameter values suggest that there may be important quantitative differences between the asymptotic second and third best tax rates.

Appendices

The competitive equilibrium

The first order condition of the agent gives:

$$c_{t+1} = c_t(\beta R(\tau_{t+1}))^\sigma \tag{0.34}$$

Iteration of the feasibility condition $k_{t+1} = \phi(\tau_t)k_t - c_t$ implies

$$c_0 + \sum_{t=1}^{T} c_t \prod_{s=1}^{t} \phi^{-1}(\tau_s) + \prod_{s=1}^{T} \phi^{-1}(\tau_s) k_{T+1} = \phi(\tau_0)k_0. \tag{0.35}$$

As is conventional we assume that

$$\lim_{T\to\infty} \prod_{s=1}^{T} \phi^{-1}(\tau_s) k_{T+1} = 0 \tag{0.36}$$

so that 0.35 becomes

$$c_0 + \sum_{t=1}^{\infty} c_t \prod_{s=1}^{t} \phi^{-1}(\tau_s) = \phi(\tau_0)k_0 \tag{0.37}$$

Iterating the first order conditions for the agent we get

$$c_t = c_0 \prod_{s=1}^{t} (\beta R(\tau_s))^\sigma \tag{0.38}$$

which, substituted into 0.37 gives

$$c_0 \left(1 + \sum_{t=1}^{\infty} \prod_{s=1}^{t} (\beta R(\tau_s))^\sigma \phi^{-1}(\tau_s) \right) = \phi(\tau_0)k_0 \tag{0.39}$$

If we now substitute the 0.38 into the expression for the utility of the agent we get that the utility from an initial capital k_0 and a tax rate sequence $\tau = (\tau_0, \tau_1, \ldots)$, denoted by $V(k_0, \tau)$, is

$$V(k_0, \tau) = (\frac{\sigma}{\sigma - 1}) c_0^{\frac{\sigma-1}{\sigma}} \left(1 + \sum_{t=1}^{\infty} \prod_{s=1}^{t} (\beta^\sigma \left(R(\tau_s) \right)^{\sigma-1}) \right) \tag{0.40}$$

Now we can use the equation 0.39 to substitute for c_0, and obtain the value to the agent in terms of k_0 and τ only. To lighten notation, we introduce:

$$X(\tau_1, \tau_2, \ldots) \equiv \left(1 + \sum_{t=1}^{\infty} \prod_{s=1}^{t} (\beta R(\tau_s))^{\sigma} \phi^{-1}(\tau_s)\right)^{\frac{1-\sigma}{\sigma}} \tag{0.41}$$

and

$$Y(\tau_1, \tau_2, \ldots) \equiv \left(1 + \sum_{t=1}^{\infty} \prod_{s=1}^{t} (\beta^{\sigma} \left(R(\tau_s)\right)^{\sigma-1})\right) \tag{0.42}$$

so that we define

$$h(\tau_1, \tau_2, \ldots) \equiv X(\tau_1, \tau_2, \ldots) Y(\tau_1, \tau_2, \ldots) \tag{0.43}$$

Now we can write

$$V(k_0, \tau) = (\frac{\sigma}{\sigma - 1}) \left(k_0\right)^{\frac{\sigma-1}{\sigma}} \left(\phi(\tau_0)\right)^{\frac{\sigma-1}{\sigma}} h\left(\tau_1, \tau_2, \ldots\right). \tag{0.44}$$

as we have in the text.

The optimal tax with commitment

Here we derive the first order conditions for the optimization. The derivative with respect to the tax rate τ_m of the two terms X and Y are as follows.

$$\begin{aligned} \frac{\partial X}{\partial \tau_m} &= \left(\frac{1-\sigma}{\sigma}\right) X^{\frac{1-2\sigma}{1-\sigma}} \sum_{t=m}^{\infty} \left(\prod_{s=1, s\neq m}^{t} (\beta R(\tau_s))^{\sigma} \phi^{-1}(\tau_s)\right) \\ &\quad \times \left(\beta \frac{\sigma \phi(\tau_m) \left(R(\tau_m)\right)^{\sigma-1} R'(\tau_m) - \phi'(\tau_m) \left(R(\tau_m)^{\sigma}\right)}{\left(\phi(\tau_m)\right)^2} \right) \end{aligned} \tag{0.45}$$

or

$$\frac{\partial X}{\partial \tau_m} = \left(\frac{1-\sigma}{\sigma}\right) X^{\frac{1-2\sigma}{1-\sigma}} \sum_{t=m}^{\infty} \left((\prod_{s=1}^{t} (\beta R(\tau_s))^{\sigma} \phi^{-1}(\tau_s)\right) \left(\sigma \frac{R'(\tau_m)}{R(\tau_m)} - \frac{\phi'(\tau_m)}{\phi(\tau_m)} \right) \tag{0.46}$$

For the Y term we get

$$\frac{\partial Y}{\partial \tau_m} = \sum_{t=m}^{\infty} \prod_{s=1, s\neq m}^{t} (\beta^{\sigma} R(\tau_s)^{\sigma-1})(\sigma - 1)(R(\tau_m)^{\sigma-2} R'(\tau_m) \tag{0.47}$$

or

$$\frac{\partial Y}{\partial \tau_m} = \sum_{t=m}^{\infty} \prod_{s=1}^{t} (\beta^{\sigma} R(\tau_s)^{\sigma-1})(\sigma - 1) \left(\frac{R'(\tau_m)}{R(\tau_m)}\right) \tag{0.48}$$

We can now substitute in the equation giving the value to the agent, to get:

$$\begin{aligned}\frac{\partial V(k_0,\tau)}{\partial \tau_m} &= (\frac{\sigma}{\sigma-1})\,(\phi(\tau_0)k_0)^{\frac{\sigma-1}{\sigma}} \\ &\times \left(Y\left(\frac{1-\sigma}{\sigma}\right)X^{\frac{1-2\sigma}{1-\sigma}}\sum_{t=m}^{\infty}\left(\prod_{s=1}^{t}(\beta R(\tau_s))^{\sigma}\phi^{-1}(\tau_s)\right)\left(\sigma\frac{R'(\tau_m)}{R(\tau_m)}-\frac{\phi'(\tau_m)}{\phi(\tau_m)}\right)\right) \\ &+(\frac{\sigma}{\sigma-1})\,(\phi(\tau_0)k_0)^{\frac{\sigma-1}{\sigma}}\left(X\sum_{t=m}^{\infty}\prod_{s=1}^{t}(\beta^{\sigma}R(\tau_s)^{\sigma-1})(\sigma-1)\frac{R'(\tau_m)}{R(\tau_m)}\right) \end{aligned} \tag{0.49}$$

Now we assume that $\beta^{\sigma}R(\tau_s)^{\sigma}\phi^{-1}(\tau_s)<1$ (it can be shown this condition is necessary for the value of the program to be bounded and for an optimum to exist). Therefore,

$$\lim_{m\to\infty}\sum_{t=m}^{\infty}\left(\prod_{s=1}^{t}(\beta R(\tau_s))^{\sigma}\phi^{-1}(\tau_s)\right)=0 \tag{0.50}$$

so that

$$\lim_{m\to\infty}\frac{\partial V(k_0,\tau)}{\partial \tau_m}=\left((\frac{\sigma}{\sigma-1})\,(\phi(\tau_0)k_0)^{\frac{\sigma-1}{\sigma}}\right)\left(X\sum_{t=m}^{\infty}\prod_{s=1}^{t}(\beta^{\sigma}R(\tau_s)^{\sigma-1})(\sigma-1)\frac{R'(\tau_m)}{R(\tau_m)}\right) \tag{0.51}$$

This, together with the optimality condition

$$\frac{\partial V(k_0,\tau)}{\partial \tau_m}=0 \tag{0.52}$$

implies that

$$\lim_{m\to\infty}R'(\tau_m)=0, \tag{0.53}$$

as claimed.

Bad Equilibria in the Barro Model

This section of the Appendix considers alternative "bad" equilibria in the Barro model.

Bad Equilibrium 1

Consider the following strategy profile. The strategies are: the government sets $\tau_t=1$ in every period, irrespective of the history. Each agent i sets the saving $k_t^i=0$ in every period, also irrespective of the history.

This is a Nash equilibrium: if the government chooses $\tau_t = 1$ in every period, then each agent will save zero, and if each agent saves zero then the government may as well tax the entire savings.

This is also a sustainable equilibrium in the sense of Chari and Kehoe [8]. The policy plan of the government is to set the tax rate equal to one after any history of tax rates, and the allocation rule for the public is to have zero savings in every period, also after every history of taxes. In this case (differently from the case of Nash equilibrium above) we have to check that policy plan and allocation rule are optimal after any history of taxes. So consider the case where $\tau_t \neq 1$. The public still expects future tax rates to be equal to one, so it will not save even after a deviation by the government to tax rates less than one. And, once again, if each agent saves zero then the government may as well tax the entire savings.

The equilibrium, however, is not subgame perfect. Consider the history where at period t the agents have saved in aggregate a positive amount (this is an off-equilibrium path, of course). Now apply the one-period-deviation principle. The government has the following problem: choose the best tax rate, considering that in the next periods (and forever) agents will save zero and the government will set the tax rate to one in future periods. The best choice is to set the tax rate to maximize present output, and that tax rate is less than one. We conclude these strategies do not give rise to a subgame perfect equilibrium.

Bad equilibrium 2

Now suppose the government chooses a tax rate equal to one in any period in which the aggregate savings are equal to zero; it chooses $\bar{\tau}$, the output-maximizing rate, in every period in which aggregate savings are positive. The agents choose savings equal to zero in every period, independently of the history of taxes, and of the value of aggregate savings they observe.

This is a subgame perfect equilibrium, and the equilibrium outcome path has zero savings and zero output in every period. To check that this is indeed a subgame perfect equilibrium, consider first any history that a single agent i is facing. The strategy we have described prescribes zero savings for her. We have to prove that this is indeed optimal for her. Consider what tax rates she is expecting. She knows that all the other agents will save zero, so that the aggregate savings rate will be zero in all future periods, and therefore (according to the postulated strategy) the tax rate of the government will be equal to one. So a deviation to positive savings by her alone will still yield zero aggregate savings –in which case the government will still choose a tax rate equal to one, and her income in the next period will still be zero. Hence, a deviation is not profitable.

The check for the optimality of government behavior is similar to the previous case. Take any history the government is facing. When aggregate savings are zero, a one-period deviation to a tax rate less than one will not give any public good, and will leave the payoff unchanged. When aggregate savings are positive, the best choice is $\bar{\tau}$. This is because the government knows that in future periods the savings will again be zero (this is what the strategy prescribes for the public), so the continuation value is constant, and independent of his choice of tax rate today. So the government wants to maximize consumption today; given that the public will save zero, this is equivalent to maximizing output, so the government will choose $\bar{\tau}$. We conclude these strategies do indeed give rise to a subgame perfect equilibrium.

However, this equilibrium is not perfect. To prove this claim, suppose there is a sequence of ϵ, tending to zero, where for each ϵ the strategy for the public is zero savings with probability $1-\epsilon$, and a random draw otherwise. For the government the strategy is to choose τ in period t as follows:

i. With probability ϵ the tax rate τ is set equal to one if the aggregate saving is zero, and equal to $\bar{\tau}$ if the aggregate saving is positive.

ii. With probability $1-\epsilon$ the tax rate τ is a random draw.

The key points to consider are the following:

(i) With probability 1 the government faces positive aggregate savings in each period, no matter what the strategy of the agents is. This follows from the assumption that strategies are totally mixed, and from the law of large numbers for a continuum of random variables.

(ii) The members of the public know that, with probability one, they will face in each period a tax rate equal to $\bar{\tau}$. This follows from the point (ii) above, and the definition of the strategy of the government. So the agents are facing their optimal allocation problem where the production function looks as follows:

i. The capital stock k_t and the flow of public services g_t in period t is small (if all the other players are following the strategy), since savings come only out of the ϵ-size mistakes;

ii. The tax rate is $\bar{\tau}$, so the return to private savings is $AR(\bar{\tau}) > 0$.

We conclude that each agent deviates and saves a positive amount. The equilibrium is therefore not perfect.

Comparison of the Chari-Kehoe setup and ours

We saw above that in the Barro model the autarchic equilibrium in which the government sets the tax rate at one and the public saves nothing is not a perfect equilibrium. In the Chari-Kehoe model, by contrast, the autarchic equilibrium in not only subgame perfect, but it can also easily be shown to be robust in the sense of being perfect.

Consider the following sketch of a proof. In this economy the payoff to each player is determined as follows. In the second period of their lives agents have income equal to $(1-\delta_t)Rk_t + (1-\tau_t)l_t$, where δ_t is the tax rate on capital, and τ_t the tax rate on labor (the rest of the notation is self-explanatory). The *autarchy equilibrium* has $\delta_t = 1$ for any period, while the tax rate τ_t on labor set optimally to provide for the tax revenues to pay the exogenous amount g.

Consider first the decision of the government. If there is no positive saving in the economy (the aggregate capital is zero), then the tax rate in the present period does not affect the current payoff, so it does not produce any one-period gain. If there are positive savings, given that agents will save almost zero in future periods the best tax rate is the maximum rate on capital, with the maximum probability $1-\epsilon$.

Consider now the agents. Expecting the maximum tax rate with high probability in the future, they calculate that a positive saving today will certainly reduce their consumption today, and increase their income tomorrow (in the second period) only in the low probability event of a zero tax rate; hence when the probability ϵ is small enough it is optimal for them to save zero with probability $1-\epsilon$, as claimed.

What explains this difference between the Chari Kehoe setup and ours? Consider the one shot problems, which are obtained by setting the discount rate equal to zero. In Chari-Kehoe the optimal tax rate on capital is one whenever there is positive capital in the economy. In our setup the optimal tax rate in the one shot problem is $\bar{\tau}$ for any initial capital which is strictly positive, and any tax rate otherwise. This explains why a deviation of the public is not important in their model, but it is in ours. Facing a positive saving, the government in their model still wants to tax capital at a maximum rate, while in our model the government wants to tax at the "myopic" rate $\bar{\tau}$.

Value of deviation

If agents expect that after a period t the tax rate $\bar{\tau}$ will be implemented, the Euler equation governing the growth of individual consumption is

$$\frac{c_{t+1}}{c_t} = (\beta R(\bar{\tau}))^{\sigma} \tag{0.54}$$

As a result,

$$c_s = c_t \left(\beta R(\bar{\tau})\right)^{\sigma(s-t)} \tag{0.55}$$

What is the level of consumption the period of the deviation? Equation 0.39 becomes

$$c_t = \left(1 + \sum_{v=t+1}^{\infty} \prod_{s=1}^{v} (\beta R(\bar{\tau}))^{\sigma} \phi^{-1}(\tau_s)\right)^{-1} \phi(\bar{\tau}) k_t \tag{0.56}$$

Therefore, the value as of period t is

$$\begin{aligned} V^D(k_t) &= \left(\frac{\sigma}{\sigma - 1}\right) \sum_{s=t}^{\infty} c_s^{\frac{\sigma-1}{\sigma}} \beta^{s-t} \\ &= \left(\frac{\sigma}{\sigma - 1}\right) (\phi(\bar{\tau}) k_t)^{\frac{\sigma-1}{\sigma}} \left(1 + \sum_{v=t+1}^{\infty} \prod_{s=1}^{v} (\beta R(\bar{\tau}))^{\sigma} \phi^{-1}(\tau_s)\right)^{\frac{1-\sigma}{\sigma}} \left(1 - \beta^{\sigma} R(\bar{\tau})^{\sigma-1}\right) \end{aligned} \tag{0.57}$$

The government must therefore select a sequence of τ's to maximize 0.57. The relevant derivative is

$$\begin{aligned} \frac{\partial V^D(k_t)}{\partial \tau_m} &= (\phi(\bar{\tau}) k_t)^{\frac{\sigma-1}{\sigma}} \left(1 + \sum_{v=t+1}^{\infty} \prod_{s=1}^{v} (\beta R(\bar{\tau}))^{\sigma} \phi^{-1}(\tau_s)\right)^{\frac{1-2\sigma}{\sigma}} \left(1 - \beta^{\sigma} R(\bar{\tau})^{\sigma-1}\right) \\ &\quad \times \sum_{v=m}^{\infty} \left(\prod_{s=1, s\neq m}^{v} (\beta R(\bar{\tau}))^{\sigma} \phi^{-1}(\tau_s)\right) \left(\beta \frac{\phi'(\tau_m)\left(R(\bar{\tau})^{\sigma}\right)}{\left(\phi(\tau_m)\right)^2}\right) \end{aligned} \tag{0.58}$$

Therefore, the first order condition for the government's problem is

$$\phi'(\tau_m) = 0 \text{ if } m > t \tag{0.59}$$

Hence, the government's best response to the individual's expectations is indeed to make them self-fulfilling by setting $\tau_m = \bar{\tau}$ if $m > t$. Since the deviation also consisted of setting $\tau_t = \bar{\tau}$, we have that $\tau_m = \bar{\tau}$ if $m \geq t$.

Using this in the expression 0.57 we have

$$\begin{aligned} V^D(k_t) &= \left(\frac{\sigma}{\sigma - 1}\right) (k_t)^{\frac{\sigma-1}{\sigma}} \left(\phi(\bar{\tau}) - (\beta R(\bar{\tau}))^{\sigma}\right)^{\frac{\sigma-1}{\sigma}} \left(1 - \beta^{\sigma} R(\bar{\tau})^{\sigma-1}\right)^{-1} \\ &= \left(\frac{\sigma}{\sigma - 1}\right) (k_t)^{\frac{\sigma-1}{\sigma}} H^*(\bar{\tau}) \end{aligned} \tag{0.60}$$

which is the equation we have in the text.

Analysis of the Lagrangean

We denote for convenience

$$H_m \equiv \frac{\partial H}{\partial \tau_m}; X_m \equiv \frac{\partial X}{\partial \tau_m}; Y_m \equiv \frac{\partial Y}{\partial \tau_m} \tag{0.61}$$

so that for any $\tau = (\tau_0, \tau_1, \ldots)$ and any $m \geq 0$,

$$\frac{\partial L}{\partial \tau_m} = H_m(\tau_0, \tau_1, \ldots) + \sum_{i=0}^{\infty} \lambda_i H_{m+1-i}(\tau_i, \tau_{i+1}, \ldots) \tag{0.62}$$

where the λ's are the Lagrange multipliers. The necessary condition for the (third best) optimality is

$$\frac{\partial L}{\partial \tau_m} = 0 \text{ for every } m \geq 0. \tag{0.63}$$

But notice:

$$H_0(\tau) = \left(\frac{\sigma - 1}{\sigma}\right) \phi\,(\tau)^{-\left(\frac{1}{\sigma}\right)} \phi'\,(\tau)\, XY \tag{0.64}$$

and

$$H_m(\tau) = \phi^{\left(\frac{\sigma-1}{\sigma}\right)} \left[X_m Y + XY_m\right] \tag{0.65}$$

Let us now consider the expression for the Lagrangean when τ_t is equal to a constant for every t. We drop the argument τ for simplicity of notation, and we find that

$$\frac{\partial X}{\partial \tau_m} = \left(\frac{1-\sigma}{\sigma}\right) X^{\left(\frac{1-2\sigma}{1-\sigma}\right)} \left(\sigma \frac{R'}{R} - \frac{\phi'}{\phi}\right) \left((\beta R)^{\sigma} \phi^{-1}\right)^{m} \left(1 - \left((\beta R)^{\sigma} \phi^{-1}\right)\right)^{-1} \tag{0.66}$$

while the other term is

$$\frac{\partial Y}{\partial \tau_m} = (\sigma - 1) \frac{R'}{R} \left[(\beta R)^{\sigma} R^{-1}\right]^{m} Y$$

If we substitute these into 0.63 and factor out the common term

$$\left(\frac{\sigma - 1}{\sigma}\right) \phi^{\left(\frac{\sigma-1}{\sigma}\right)} XY$$

we find that 0.63 is equivalent to

$$-\left(\sigma \frac{R'}{R} - \frac{\phi'}{\phi}\right) \sum_{i=0}^{m-1} \lambda_i \left((\beta R)^{\sigma} \phi^{-1}\right)^{m+1-i} + \frac{\sigma R'}{R} \sum_{i=0}^{m-1} \lambda_i \left((\beta R)^{\sigma} R^{-1}\right)^{i} + \lambda_m \frac{\phi'}{\phi} = 0$$

If we substitute an exponential λ solution of the above equation, with $\lambda_i = \lambda^i$, we obtain

$$-\left(\sigma \frac{R'}{R} - \frac{\phi'}{\phi}\right) \frac{\lambda^m - \left((\beta R)^{\sigma} \phi^{-1}\right)^{m-1}}{\lambda - \left((\beta R)^{\sigma} \phi^{-1}\right)} + \frac{\sigma R'}{R} \frac{\lambda^m - \left((\beta R)^{\sigma} R^{-1}\right)^{m-1}}{\lambda - \left((\beta R)^{\sigma} R^{-1}\right)} + \lambda^m \frac{\phi'}{\phi} = 0. \tag{0.67}$$

It is easy to see that the only exponential solution of the above equation must satisfy the condition on λ:

$$1 > \lambda > \left((\beta R)^{\sigma} R^{-1}\right) \tag{0.68}$$

The condition $1 > \lambda$, in particular, follows from the fact that the Lagrange multipliers have to be summable. In addition, if the condition 0.68 is satisfied, then in the limit as $m \to \infty$, λ^m dominates $\left((\beta R)^{\sigma}\phi^{-1}\right)^{m-1}$ and the equation 0.67 reduces to

$$-\left(\sigma\frac{R'}{R}-\frac{\phi'}{\phi}\right)[\lambda-\left((\beta R)^{\sigma}\phi^{-1}\right)]^{-1}+\sigma\frac{R'}{R}[\lambda-\left((\beta R)^{\sigma}R^{-1}\right)]^{-1}+\frac{\phi'}{\phi}=0. \tag{0.69}$$

We have now to check that there exists a solution in λ of the above equation, where λ satisfies the additional condition 0.68. Let us define the function $\Phi(\lambda)$ as

$$\begin{aligned}\Phi(\lambda) \quad &\equiv \quad -(\sigma\frac{R'}{R}-\frac{\phi'}{\phi})(\lambda-\frac{(\beta R)^{\sigma}}{R})+\sigma\frac{R'}{R}(\lambda-\frac{(\beta R)^{\sigma}}{\phi}) \\ &\quad +\frac{\phi'}{\phi}(\lambda-\frac{(\beta R)^{\sigma}}{R})(\lambda-\frac{(\beta R)^{\sigma}}{\phi})\end{aligned} \tag{0.70}$$

Note that λ^* solves 0.69 if and only if $\Phi(\lambda^*)=0$.

In what follows we focus only on the case of $\rho < 0$, so that $\bar{\tau} > \tilde{\tau}$. An analogous proof exists for the converse case. If $\rho < 0$. We prove that a λ as required exists by observing that Φ is a continuous function of λ and that

$$\Phi(\frac{(\beta R)^{\sigma}}{R}) < 0; \Phi(1) > 0. \tag{0.71}$$

The first inequality follows from direct substitution, which gives:

$$\Phi(\frac{(\beta R)^{\sigma}}{R}) = (\frac{(\beta R)^{\sigma}}{R})(\sigma R')(\frac{1}{R}-\frac{1}{\phi})$$

a non-positive number since $R' \leq 0$ and $R \leq \phi$; for the second, we have:

$$\begin{aligned}\Phi(1) \quad &= \quad -(\frac{\sigma R'}{R}-\frac{\phi'}{\phi})[1-(\frac{(\beta R)^{\sigma}}{\phi})]^{-1}+ \\ &\quad (\frac{\sigma R'}{R})[1-(\frac{(\beta R)^{\sigma}}{R})]^{-1}+\frac{\phi'}{\phi}.\end{aligned}$$

The derivative of H^* is now:

$$\frac{dH^*}{d\tau} \equiv -(\frac{\sigma R'}{R} - \frac{\phi'}{\phi})(\frac{(\beta R)^\sigma}{\phi})[1-(\frac{(\beta R)^\sigma}{\phi})]^{-1} + (\frac{\sigma R'}{R})(\frac{(\beta R)^\sigma}{R})[1-(\frac{(\beta R)^\sigma}{R})]^{-1} + \frac{\phi'}{\phi}.$$

Now we take the difference:

$$\Phi(1) - \frac{dH^*}{d\tau} = \frac{\phi'}{\phi}$$

which is positive, because both ϕ and ϕ' are positive; also $\frac{dH^*}{d\tau}$ is positive. We conclude that $\Phi(1) > 0$, and our claim is proved.

We can also show that the λ^* is unique, and that there is no λ that solves $\Phi(\lambda) = 0$ for $1 > \lambda > ((\beta R)^\sigma R^{-1})$ if the condition given by 0.71 fails. For this purpose it suffices to show that

$$\Phi'(\lambda) = \frac{\phi'}{\phi}\left(1 + 2\lambda - \frac{(\beta R)^\sigma}{R} - \frac{(\beta R)^\sigma}{\phi}\right) > 0 \tag{0.72}$$

over the interval $1 > \lambda > ((\beta R)^\sigma R^{-1})$. Notice however that $1 + 2\lambda - \frac{(\beta R)^\sigma}{R} - \frac{(\beta R)^\sigma}{\phi} > 0$ because $\lambda > (\beta R)^\sigma R^{-1} > (\beta R)^\sigma \phi^{-1}$. Therefore sign $\Phi'(\lambda) = \text{sign}\frac{\phi'}{\phi} > 0$, since we are in the case where $\rho < 0$. Now it is clear by inspection that if 0.71 fails there is no λ^* such that $\Phi(\lambda^*) = 0$ for $1 > \lambda^* > ((\beta R)^\sigma R^{-1})$. The argument is the same for $\rho > 0$ except for appropriate modifications of signs and inequalities.

References

[1] Barro, R. J.; (1990). Government Spending in a Simple Model of Endogenous Growth. *Journal of Political Economy*, **98, 5**, part **2**, S103-S125.

[2] Barro, R. J. and. Gordon D. B., Rules, Discretion and Reputation in a Model of Monetary Policy. *Journal of Monetary Economics,* **12,** 101-122.

[3] Benhabib, J. and Rustichini, A.; (1996). Social Conflict and Growth. *Journal of Economic Growth.* **1**, 125-142.

[4] Benhabib, J. and Velasco, A.; (1996). On the Optimal and Best Sustainable Taxes in an Open Economy. *European Economic Review,* **40**, 135-154.

[5] Benhabib, J. and Rustichini, A. ; (1995). Optimal Taxation without Committment, CORE DP.

[6] Chamley, C.; (1986). Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives. *Econometrica*, **54, 3**, 607-622.

[7] Chamley, C. ; (1985). Efficient Taxation in a Stylized Model of Intertemporal General Equilibrium. *International Economic Review*, **26, 2**, 451-468.

[8] Chari, V V. and Kehoe, P. J. ;(1990). Sustainable Plans. *Journal of Political Economy,* **98,** 783-802.

[9] Chari, V.V., Christiano L. J., and Kehoe, P.; (1994). Optimal Fiscal Policy in a Business Cycle Model. *Journal of Political Economy*, **102, 4**, 617-652.

[10] Jones, L. Manuelli, R.E. and Rossi, P.; (1993) Optimal Taxation in Models of Endogenous Growth. *Journal of Political Economy.*

[11] Judd, K. ;(1985). Redistributive Taxation in a Simple Perfect Foresight Model. *Journal of Public Economics*, **28**, 59-83.

[12] Kydland, F. E. and Prescott, E. C. : (1977). Rules Rather Than Discretion: The Time Consistency of Optimal Plans. *Journal of Political Economy,* **85,** 473-491.

[13] Lucas, R. E. and Stokey, N.; (1983) Optimal Fiscal and Monetary Policy in an Economy without Capital. *Journal of Monetary Economics,* **12**, 55-93.

[14] Marcet, A., and Marimon, R.; (1994). Recursive Contracts. working paper, Universitat Pompeu Fabra, Barcelona.

[15] Ramsey, F. P.; (1928). A Mathematical Theory of Saving. *Economic Journal*, **38**, 543-559.

[16] Stokey, N. L.; (1991). Credible Public Policy. *Journal of Economic Dynamics and Control,* **15**, 627-657.