ECONOMIC RESEARCH REPORTS

PUBLIC CAPITAL AND OPTIMAL TAXES WITHOUT COMMITMENT

by **Jess Benhabib,**
Aldo Rustichini,
and
Andrés Velasco

RR # 96-19 **May 1996**

C.V. STARR CENTER FOR APPLIED ECONOMICS

NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, NY 10003-6687

Public Capital and Optimal Taxes Without Commitment*

Jess Benhabib
New York University

Aldo Rustichini
CORE

Andrés Velasco
New York University and NBER

May 6, 1996

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.

JEL Classification: H21, D9
Key Words: Public Capital, Optimal Taxation, Time Consistency

---

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

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 capital. 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. 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 differences in tax rates between the second and third best policies for a wide range of parameters. It turns out that these differences in the tax rates can be quite significant.

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].[1] 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 public goods.[2] This is true of in a broad class of intertemporal tax problems.

---

[1] 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.

[2] 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,

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.

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 capital that maximize single-period output). Since agents would expect this and save accordingly, the outcome would be a bad (discretionary) equilibrium with low savings.[3] 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. 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.

But it is also possible –especially in a model with capital accumulation– 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 capital in circumstances where a commitment by the government to future policies is not possible.

In the model developed by Barro [1] there are no fixed factors. Therefore, and unlike the models of Chamley [6], [7] and Judd [11], accumulation leads to persistent endogenous growth. Barro uses a Cobb-Douglas production function for output, with constant return to scale and private and public capital 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

---

is presented in Benhabib and Rustichini [3].

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

longer constant (in particular, the initial-period tax rate is different from that of all subsequent periods), and the time-consistency problem emerges. We first 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 specification leading to endogenous growth, may be larger or smaller that the output-maximizing rate, depending on whether the elasticity of substitution between private and public capital is above or below one. [4]

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 the best sustainable asymptotic tax rates must be 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. 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, requiring tax rates to be time consistent leads to lower economic growth.

The next section describes the model and the equilibrium. Section 3 discusses the optimal taxes under commitment. Section 4 describes the value of deviation. Section 5 sets up the problem with incentive constraints when

---

[4] 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.

commitment by the government to future tax rate is not possible. Section 5 provides some particular examples and provides a characterization of the optimal taxes without commitment. Finally section 6 provides a family of calibrated examples to illustrate numerically the differences between optimal taxes with and without commitment.

2 An Economy with Private and Public Capital

We first describe our simple economy. There is one representative agent, who has an infinite life, and two goods. One is a private good, which is used both in production and in consumption. The other is a public capital good which is only used in production. 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 and public capital 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 amount of the public capital good, $\rho \in [-1, +\infty]$, $A \geq 0$. 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 capital 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 capital.

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):

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

Where clearly $R(\tau) > 0$ for every τ.[5] Note that

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

by the homogeneity of degree one of $\phi(\tau)$ 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 $R/\phi = a$.

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:

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

---

[5] If one wanted to ensure that clearly $R(\tau) > 1$ 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 $R(\tau) > 1$ for all allowable tax rates.

The equilibrium

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 is given by a function $V(k_0, \tau)$, which is found to have the simple form

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

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

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

When the terms of the sequence τ are constant over time then H has a simple form

$$H(\tau, \tau, \ldots) = (\phi\,(\tau) - (\beta R(\tau))^\sigma)^{\frac{\sigma-1}{\sigma}} \left(1 - \beta^\sigma R(\tau)^{\sigma-1}\right)^{-1} \tag{2.11}$$

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

$$H^*(\tau) \equiv H(\tau, \tau, \ldots). \tag{2.12}$$

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 capital, 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

$$\lim_{t \to \infty} R'(\tau_t) = 0 \tag{3.13}$$

where $\tilde{\tau}$ is henceforth defined as the one that solves $R'(\tau_t) = 0$. 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. Under the assumed CES technology, the tax rate given by 3.13 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 the condition 3.13 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 the case of Cobb-Douglas technology ($\rho = 0$) studied by Barro [1], $R(\tau)/\phi(\tau) = a$, so that $\tilde{\tau} = \bar{\tau}$. 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 $\tilde{\tau} = \bar{\tau}$.

We now consider in detail two examples of economies that will probably help to develop intuition about the way the limit of second best tax rates is determined and, later on, about the way the incentive compatibility constraint operates.

Logarithmic utility, linear production

With $\sigma = 1$ (or, more precisely, when the utility function is $\log c_t$), and $\rho = -1$, $A = 1$, (*i.e.* the production function is linear), we have

$$\phi(\tau) = a(1-\tau) + (1-a)\tau, \text{ and } R(\tau) = a(1-\tau) \tag{3.14}$$

In the Appendix the reader will find more detailed computation for the case of logarithmic utility; there we find that

$$H^*(\tau) = (1-\beta)^{-1}\log(\phi(\tau) - \beta R(\tau)) + \beta(1-\beta)^{-2}\log R(\tau) + \text{ constant} \tag{3.15}$$

It follows that

$$\bar{\tau} = 1 \text{ if } a < 1/2, 0 \text{ if } a > 1/2; \tilde{\tau} = 0 \text{ for every } a \tag{3.16}$$

So if $a > 1/2$, private capital is relatively more productive, and the optimal tax is a zero tax. The second best is trivially incentive compatible. On the other hand, if $a < 1/2$, then there is a genuine trade-off: public capital is always more productive than private capital, but tax rates have to be kept below 1 to stimulate private accumulation.

In this case, the condition that the limit tax rate with commitment should maximize the private rate of return R gives that along the tax sequence $\{\tau_t\}$:

$$\lim_{t\to\infty} \tau_t = 0 \tag{3.17}$$

It is interesting to note that the zero tax rate in 3.17 is in fact achieved only in the limit; *i.e.* that $\tau_t > 0$ for every $t > 0$. The proof of the claim –to be found in the Appendix– is based on the analysis of the appropriate first order conditions.

Logarithmic utility, Leontief technology

At the opposite extreme of the linear production function we have the case where $\rho = \infty$, so that

$$\phi(\tau) = A \min\{a(1-\tau), \tau(1-a)\} \tag{3.18}$$

and

$$R(\tau) = Aa(1-\tau) \text{ if } a(1-\tau) \leq (1-a)\tau; \text{and } = 0 \text{ otherwise.} \tag{3.19}$$

The formula $R'(\tilde{\tau}) = 0$ can no longer be directly applied, although the conjecture $\tilde{\tau} = 0$ is natural; one can in fact prove that this is the limit tax rate in the commitment case.

4 The value of deviation

We now consider the possibility of reputational equilibria. We assume that 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.

Recall $\bar{\tau}$ is the solution of maximizing $\phi(\tau)$, that is $\phi'(\bar{\tau}) = 0$. 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

$$V^D(k_t) = \left(\frac{\sigma}{\sigma-1}\right)(k_t)^{\frac{\sigma-1}{\sigma}} H^*(\bar{\tau}) \tag{4.20}$$

where

$$H^*(\bar{\tau}) \equiv \left(\phi(\bar{\tau}) - (\beta R(\bar{\tau}))^{\sigma}\right)^{\frac{\sigma-1}{\sigma}} \left(1 - \beta^{\sigma} R(\bar{\tau})^{\sigma-1}\right)^{-1} \tag{4.21}$$

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, $V^C(k_t, \tau)$

$$V^C(k_t, \tau) = \left(\frac{\sigma}{\sigma - 1}\right) (\phi(\tau_t)k_t)^{\frac{-1}{\sigma^\sigma}} h\left(\tau_{t+1}, \tau_{t+2}, ...\right) \tag{4.22}$$

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

$$V^C\left(k_t, \tau\right) - V^D(k_t) = \left(\frac{\sigma}{\sigma - 1}\right) k_t^{\frac{\sigma-1}{\sigma}} \left((\phi(\tau_t))^{\frac{-1}{\sigma^\sigma}} h\left(\tau_{t+1}, \tau_{t+2}, ...\right) - H^*\left(\bar{\tau}\right)\right) \tag{4.23}$$

is non negative if and only if

$$(\phi(\tau_t))^{\frac{\sigma-1}{\sigma}} h\left(\tau_{t+1}, \tau_{t+2}, ...\right) - H^*\left(\bar{\tau}\right) \geq 0, \text{ for all } t \geq 1 \tag{4.24}$$

The remarkable feature of 4.24 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 4.24 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 4.24 reduce to a single constraint. In the general case, however, we have to take into account the entire (infinite) set.

5 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 4.24, 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

$$\max_{\tau=(\tau_0,\tau_1,\ldots)} V(k_0,\tau) \tag{5.25}$$

$$\text{subject to } V^C(k_t,\tau) \geq V^D(k_t) \text{ for all } t \geq 1 \tag{5.26}$$

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

$$\max_{\tau=(\tau_0,\tau_1,\ldots)} \left(\frac{\sigma}{\sigma-1}\right) H(\tau_0,\tau_1,\ldots) \tag{5.27}$$

$$\text{subject to } \left(\frac{\sigma}{\sigma-1}\right) H(\tau_t,\tau_{t+1},\ldots) \geq \left(\frac{\sigma}{\sigma-1}\right) H^*(\tau_p) \text{ for all } t \geq 1. \tag{5.28}$$

This problem has a Lagrangean:

$$L(\tau,\lambda) \equiv H(\tau_0,\tau_1,\ldots) + \sum_{t=1}^{\infty} \lambda_t \left[H(\tau_t,\tau_{t+1},\ldots) - H^*(\tau_p)\right]. \tag{5.29}$$

where the λ's are the Lagrange multipliers.

Existence of the optimal tax in this third best problem is proved in the Appendix. The necessary condition for the (third best) interior solutions is

$$\frac{\partial L}{\partial \tau_m} = 0 \text{ for every } m \geq 0. \tag{5.30}$$

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:

$$R'(\tilde{\tau}) = 0. \tag{5.31}$$

The tax plan with such a limit tax rate may or may not be incentive compatible. 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

$$H^*(\tilde{\tau}) - H^*(\bar{\tau}) \tag{5.32}$$

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. Our two basic examples, linear technology and Leontief technology, may help to understand how this constraint on government policy operates.

Logarithmic utility, linear production.

We let $a < 1/2$ to concentrate attention on the interesting case. Otherwise, since the production function is linear, there is no reason to have any public capital. The value of defection is somewhat degenerate in this case, since $\bar{\tau} = 1$, and no private capital is accumulated in any of the periods after deviation, so the value of deviation is $-\infty$.

To make the problem more interesting, we assume that the tax rate τ has to be contained in the interval $[0, \tau_B]$, with $\tau_B < 1$. Then

$$\tilde{\tau} = 0, \bar{\tau} = \tau_B \tag{5.33}$$

and the difference between the optimal plan and the deviation is (up to a positive multiplicative factor):

$$H^*(\tilde{\tau}) - H^*(\bar{\tau}) = \log\left(\frac{a(1-\beta)}{(1-a)\tau_B + a(1-\tau_B)(1-\beta)}\right) - \left(\frac{\beta}{1-\beta}\right)\log(1-\tau_B) \tag{5.34}$$

It is clear that the difference is negative if and only if β is smaller than some critical value, strictly included in $(0, 1)$. So, for β smaller than this value the incentive constraint is binding and the third best solution converges to a strictly positive tax rate.

Logarithmic utility, Leontief technology.

In this case $\phi(\tau) = A\min\{a(1-\tau), \tau(1-a)\}$, and the deviation tax rate is $\bar{\tau} = a$. Therefore, from what we have seen already for the commitment optimal tax,

$$\tilde{\tau} = \bar{\tau} = a \tag{5.35}$$

The Leontief technology is an extreme case in which the incentive constraints are not important. With this technology maximizing output is equivalent to maximizing the return on private capital and also the growth rate. In other words the optimal tax policy is the same as the defection policy.

In section 6 we shall see how this result will be approximated for the cases where the value of ρ is large.

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.

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

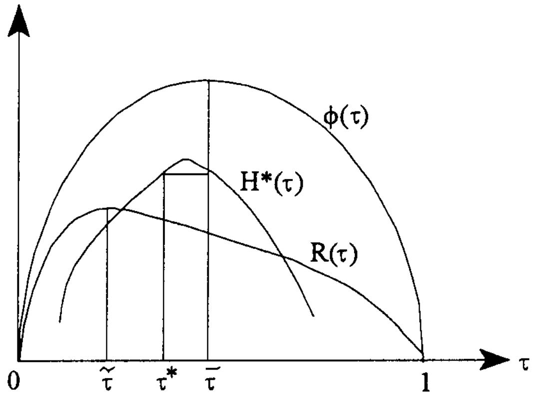

Figure 1: $\rho < 0$

$$\phi'(\tau) = \phi(\tau)^{1+\rho}[(1-a)\tau^{-(1+\rho)} - a(1-\tau)^{-(1+\rho)}]; \bar{\tau} = [1+\left(\frac{a}{1-a}\right)^{\frac{1}{1-\rho}}]^{-1} \quad (5.36)$$

and

$$R'(\bar{\tau}) = a\rho\phi(\bar{\tau})^{1+\rho}(1-\bar{\tau})^{-(1+\rho)} \quad (5.37)$$

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 $\rho < 0$.[6]

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 $[\tilde{\tau}, \bar{\tau}]$.

Now define τ^* formally to be

---

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

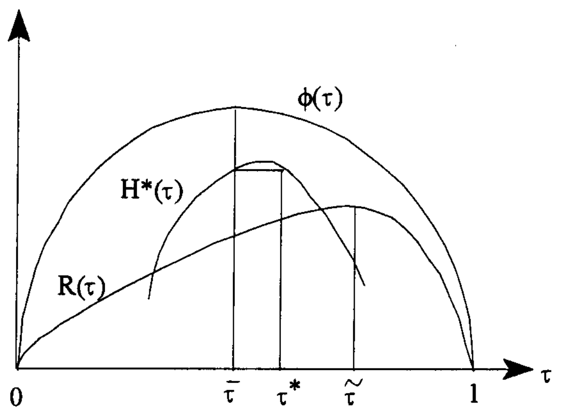

Figure 2: $\rho > 0$

$$\min\{\tau \in [\tilde{\tau}, \bar{\tau}] : H^*(\tau) = H^*(\bar{\tau})\} \tag{5.38}$$

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

$$\frac{dH^*}{d\tau}(\tau^*) \geq 0, \text{ and } \frac{dR}{d\tau}(\tau^*) \leq 0 \leq \frac{d\phi}{d\tau}(\tau^*). \tag{5.39}$$

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 8.78). Since implementing τ^H in the limit yields $H\left(\tau^H\right) > H\left(\overline{\tau}\right)$, the incentive constraint is not binding and the associated Lagrange multiplier must be zero, in contradiction to the condition 8.78 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.*

6 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 and public capital. 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.

7 Conclusions

We consider an economy where a benevolent government faces a trade-off between supplying productive public capital 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.

8 Appendices

The equilibrium

The first order condition of the agent give:

$$c_{t+1} = c_t(\beta R(\tau_{t+1}))^\sigma \tag{8.40}$$

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

$$c_0 + \sum_{t=1}^{T} c_t \prod_{s=1}^{t} \phi^{-1}(\tau_s) + \prod_{s=1}^{T} \phi^{-1}(\tau_s) k_{T+1} = \phi(\tau_0)k_0. \tag{8.41}$$

As is conventional we assume that

$$\lim_{T\to\infty} \prod_{s=1}^{T} \phi^{-1}(\tau_s) k_{T+1} = 0 \tag{8.42}$$

so that 8.41 becomes

$$c_0 + \sum_{t=1}^{\infty} c_t \prod_{s=1}^{t} \phi^{-1}(\tau_s) = \phi(\tau_0)k_0 \tag{8.43}$$

Iterating the first order conditions for the agent we get

$$c_t = c_0 \prod_{s=1}^{t} (\beta R(\tau_s))^\sigma \tag{8.44}$$

which, substituted into 8.43 gives

$$c_0 \left(1 + \sum_{t=1}^{\infty} \prod_{s=1}^{t} (\beta R(\tau_s))^\sigma \phi^{-1}(\tau_s)\right) = \phi(\tau_0)k_0 \tag{8.45}$$

If we now substitute the 8.44 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

$$V(k_0, \tau) = (\frac{\sigma}{\sigma - 1}) c_0^{\frac{\sigma-1}{\sigma}} \left(1 + \sum_{t=1}^{\infty} \prod_{s=1}^{t} (\beta^\sigma (R(\tau_s))^{\sigma-1})\right) \tag{8.46}$$

Now we can use the equation 8.45 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:

$$X(\tau_1, \tau_2, \ldots) \equiv \left(1 + \sum_{t=1}^{\infty} \prod_{s=1}^{t} (\beta R(\tau_s))^\sigma \phi^{-1}(\tau_s)\right)^{\frac{1-\sigma}{\sigma}} \tag{8.47}$$

and

$$Y(\tau_1, \tau_2, \ldots) \equiv \left(1 + \sum_{t=1}^{\infty} \prod_{s=1}^{t} (\beta^{\sigma} \left(R(\tau_s)\right)^{\sigma-1})\right) \tag{8.48}$$

so that we define

$$h(\tau_1, \tau_2, \ldots) \equiv X(\tau_1, \tau_2, \ldots) Y(\tau_1, \tau_2, \ldots) \tag{8.49}$$

Now we can write

$$V(k_0, \tau) = (\frac{\sigma}{\sigma - 1}) \, (k_0)^{\frac{-1}{\sigma^{\sigma}}} \, (\phi(\tau_0))^{\frac{-1}{\sigma^{\sigma}}} \, h\,(\tau_1, \tau_2, \ldots) \, . \tag{8.50}$$

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.

$$\begin{aligned} \frac{\partial X}{\partial \tau_m} &= \left(\frac{1-\sigma}{\sigma}\right) X^{\frac{1-2\sigma}{1-\sigma}} \sum_{t=m}^{\infty} \left(\prod_{s=1, s \neq m}^{t} (\beta R(\tau_s))^{\sigma} \phi^{-1}(\tau_s) \right) \\ & \quad \times \left(\beta \frac{\sigma \phi(\tau_m) \left(R(\tau_m)\right)^{\sigma-1} R'(\tau_m) - \phi'(\tau_m) \left(R(\tau_m)^{\sigma}\right)}{\left(\phi(\tau_m)\right)^2} \right) \end{aligned} \tag{8.51}$$

or

$$\frac{\partial X}{\partial \tau_m} = \left(\frac{1-\sigma}{\sigma}\right) X^{\frac{1-2\sigma}{1-\sigma}} \sum_{t=m}^{\infty} \left((\prod_{s=1}^{t} (\beta R(\tau_s))^{\sigma} \phi^{-1}(\tau_s) \right) \left(\sigma \frac{R'(\tau_m)}{R(\tau_m)} - \frac{\phi'(\tau_m)}{\phi(\tau_m)} \right) \tag{8.52}$$

For the Y term we get

$$\frac{\partial Y}{\partial \tau_m} = \sum_{t=m}^{\infty} \prod_{s=1, s \neq m}^{t} (\beta^{\sigma} R(\tau_s)^{\sigma-1})(\sigma - 1)(R(\tau_m)^{\sigma-2} R'(\tau_m) \tag{8.53}$$

or

$$\frac{\partial Y}{\partial \tau_m} = \sum_{t=m}^{\infty} \prod_{s=1}^{t} (\beta^{\sigma} R(\tau_s)^{\sigma-1})(\sigma - 1) \left(\frac{R'(\tau_m)}{R(\tau_m)} \right) \tag{8.54}$$

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

$$
\begin{aligned}
\frac{\partial V(k_0,\tau)}{\partial \tau_m} &= (\frac{\sigma}{\sigma-1})\,(\phi(\tau_0)k_0)^{\frac{-1}{\sigma^\sigma}} \\
&\times \left(Y\left(\frac{1-\sigma}{\sigma}\right)X^{\frac{1-2\sigma}{1-\sigma}}\sum_{t=m}^{\infty}\left(\prod_{s=1}^{t}(\beta R(\tau_s))^\sigma \phi^{-1}(\tau_s)\right)\left(\sigma\frac{R'(\tau_m)}{R(\tau_m)}-\frac{\phi'(\tau_m)}{\phi(\tau_m)}\right)\right) \\
&+(\frac{\sigma}{\sigma-1})\,(\phi(\tau_0)k_0)^{\frac{-1}{\sigma^\sigma}}\left(X\sum_{t=m}^{\infty}\prod_{s=1}^{t}(\beta^\sigma R(\tau_s)^{\sigma-1})(\sigma-1)\frac{R'(\tau_m)}{R(\tau_m)}\right) \qquad (8.55)
\end{aligned}
$$

Now we assume that $\beta^\sigma R(\tau_s)^\sigma \phi^{-1}(\tau_s) < 1$; this condition is necessary for the value of the program to be bounded and for an optimum to exist. (See the existence proof in section 8: the condition there is given as $\beta^\sigma R(\tau_s)^{\sigma-1} < 1$ which implies the previous one since $R < \phi$ for all τ.) Therefore,

$$\lim_{m\to\infty}\sum_{t=m}^{\infty}\left(\prod_{s=1}^{t}(\beta R(\tau_s))^\sigma \phi^{-1}(\tau_s)\right) = 0 \qquad (8.56)$$

so that

$$\lim_{m\to\infty}\frac{\partial V(k_0,\tau)}{\partial \tau_m} = \left((\frac{\sigma}{\sigma-1})\,(\phi(\tau_0)k_0)^{\frac{\sigma-1}{\sigma}}\right)\left(X\sum_{t=m}^{\infty}\prod_{s=1}^{t}(\beta^\sigma R(\tau_s)^{\sigma-1})(\sigma-1)\frac{R'(\tau_m)}{R(\tau_m)}\right) \qquad (8.57)$$

This, together with the optimality condition

$$\frac{\partial V(k_0,\tau)}{\partial \tau_m} = 0 \qquad (8.58)$$

implies that

$$\lim_{m\to\infty} R'(\tau_m) = 0, \qquad (8.59)$$

as claimed.

Logarithmic utility, Linear Production

For convenience we report here the expression for the function H in the case of a logarithmic utility function:

$$H(\tau_0,\ldots) = (1-\beta)^{-1}\log\left(\phi(\tau_0)\left[1+\sum_{t=1}^{\infty}\prod_{s=1}^{t}\beta\frac{R(\tau_s)}{\phi(\tau_s)}\right]^{-1}\right)+\sum_{t=1}^{\infty}\beta^t\log\prod_{s=1}^{t}\beta R(\tau_s) \qquad (8.60)$$

so that first order conditions are

$$\frac{\partial H}{\partial \tau_m} = -\frac{\left(\sum_{t=m}^{\infty}\prod_{s=1}^{t}\beta\frac{R(\tau_s)}{\phi(\tau_s)}\right)\left(\frac{R'(\tau_m)}{R(\tau_m)} - \frac{\phi'(\tau_m)}{\phi(\tau_m)}\right)}{1+\sum_{t=1}^{\infty}\prod_{s=1}^{t}\beta\frac{R(\tau_s)}{\phi(\tau_s)}} + \sum_{t=m}^{\infty}\frac{R'(\tau_m)}{\phi(\tau_m)}\beta^t, \text{ if } m \geq 1; \tag{8.61}$$

and

$$\frac{\partial H}{\partial \tau_m} = (1-\beta)^{-1}\frac{\phi'(\tau_0)}{\phi(\tau_0)}, \text{ if } m = 0. \tag{8.62}$$

Now we turn to the claim in the main text that, with log utility and linear production, the tax rate is always strictly positive and only asymptotically zero. The expression for the derivative of H that we have just derived, evaluated at $\rho = -1$ (linear production) and $\tau_m = 0$ is

$$\frac{\partial H}{\partial \tau_m} = \left(\frac{1-a}{a}\right) - 1 = \frac{1-2a}{a}, \text{ if } m \geq 1 \tag{8.63}$$

Since $\frac{1-2a}{a} > 0$ if $a < (1/2)$ (as is necessary for the problem to be non-trivial), it follows that $\tau_m = 0$ cannot be a solution for $m \geq 1$ and finite.

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

$$\frac{c_{t+1}}{c_t} = (\beta R(\bar{\tau}))^{\sigma} \tag{8.64}$$

As a result,

$$c_s = c_t\,(\beta R(\bar{\tau}))^{\sigma(s-t)} \tag{8.65}$$

What is the level of consumption the period of the deviation? Equation 8.45 becomes

$$c_t = \left(1 + \sum_{v=t+1}^{\infty}\prod_{s=1}^{v}(\beta R(\bar{\tau}))^{\sigma}\phi^{-1}(\tau_s)\right)^{-1}\phi(\bar{\tau})k_t \tag{8.66}$$

Therefore, the value as of period t is

$$\begin{aligned} V(k_t) &= \left(\frac{\sigma}{\sigma-1}\right)\sum_{s=t}^{\infty}c_s^{\frac{\sigma-1}{\sigma}}\beta^{s-t} \\ &= \left(\frac{\sigma}{\sigma-1}\right)(\phi(\bar{\tau})k_t)^{\frac{\sigma-1}{\sigma}}\left(1+\sum_{v=t+1}^{\infty}\prod_{s=1}^{v}(\beta R(\bar{\tau}))^{\sigma}\phi^{-1}(\tau_s)\right)^{\frac{1-\sigma}{\sigma}}\left(1-\beta^{\sigma}R(\bar{\tau})^{\sigma-1}\right) \end{aligned} \tag{8.67}$$

The government must therefore select a sequence of τ's to maximize 8.67. The relevant derivative is

$$\begin{aligned}\frac{\partial V^D(k_t)}{\partial \tau_m} &= (\phi(\bar{\tau})k_t)^{\frac{-1}{\sigma^\sigma}} \left(1 + \sum_{v=t+1}^{\infty} \prod_{s=1}^{v} (\beta R(\bar{\tau}))^\sigma \phi^{-1}(\tau_s)\right)^{\frac{1-2\sigma}{\sigma}} \left(1 - \beta^\sigma R(\bar{\tau})^{\sigma-1}\right) \\ &\quad \times \sum_{v=m}^{\infty} \left(\prod_{s=1, s\neq m}^{v} (\beta R(\bar{\tau}))^\sigma \phi^{-1}(\tau_s)\right) \left(\beta \frac{\phi'(\tau_m)\left(R(\bar{\tau})^\sigma\right)}{\left(\phi(\tau_m)\right)^2}\right) \qquad (8.68)\end{aligned}$$

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

$$\phi'(\tau_m) = 0 \text{ if } m > t \qquad (8.69)$$

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 8.67 we have

$$\begin{aligned}V^D(k_t) &= \left(\frac{\sigma}{\sigma-1}\right)(k_t)^{\frac{\sigma-1}{\sigma}} \left(\phi(\bar{\tau}) - (\beta R(\bar{\tau}))^\sigma\right)^{\frac{\sigma-1}{\sigma}} \left(1 - \beta^\sigma R(\bar{\tau})^{\sigma-1}\right)^{-1} \\ &= \left(\frac{\sigma}{\sigma-1}\right)(k_t)^{\frac{\sigma-1}{\sigma}} H^*(\bar{\tau}) \qquad (8.70)\end{aligned}$$

which is the equation we have in the text.

Analysis of the Lagrangean

We denote for convenience

$$H_m \equiv \frac{\partial H}{\partial \tau_m}; X_m \equiv \frac{\partial X}{\partial \tau_m}; Y_m \equiv \frac{\partial Y}{\partial \tau_m} \qquad (8.71)$$

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

$$\frac{\partial L}{\partial \tau_m} = H_m(\tau_0, \tau_1, \ldots) + \sum_{i=0}^{\infty} \lambda_i H_{m+1-i}(\tau_i, \tau_{i+1}, \ldots) \qquad (8.72)$$

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

$$\frac{\partial L}{\partial \tau_m} = 0 \text{ for every } m \geq 0. \qquad (8.73)$$

But notice:

$$H_0(\tau) = \left(\frac{\sigma-1}{\sigma}\right)\phi(\tau)^{-\left(\frac{1}{\sigma}\right)}\phi'(\tau)XY \qquad (8.74)$$

and

$$H_m(\tau) = \phi^{\left(\frac{\sigma-1}{\sigma}\right)} \left[X_m Y + X Y_m\right] \tag{8.75}$$

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

$$\frac{\partial X}{\partial \tau_m} = \left(\frac{1-\sigma}{\sigma}\right) X^{\left(\frac{1-2\sigma}{1-\sigma}\right)} \left(\sigma \frac{R'}{R} - \frac{\phi'}{\phi}\right) \left((\beta R)^\sigma \phi^{-1}\right)^m \left(1 - \left((\beta R)^\sigma \phi^{-1}\right)\right)^{-1} \tag{8.76}$$

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 8.73 and factor out the common term

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

we find that 8.73 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

$$-\left(\sigma \frac{R'}{R} - \frac{\phi'}{\phi}\right) \frac{\lambda^m - ((\beta R)^\sigma \phi^{-1})^{m-1}}{\lambda - ((\beta R)^\sigma \phi^{-1})} + \frac{\sigma R'}{R} \frac{\lambda^m - ((\beta R)^\sigma R^{-1})^{m-1}}{\lambda - ((\beta R)^\sigma R^{-1})} + \lambda^m \frac{\phi'}{\phi} = 0. \tag{8.77}$$

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

$$1 > \lambda > \left((\beta R)^\sigma R^{-1}\right) \tag{8.78}$$

The condition $1 > \lambda$, in particular, follows from the fact that the Lagrange multipliers have to be summable. In addition, if the condition 8.78 is satisfied, then in the limit as $m \to \infty$, λ^m dominates $((\beta R)^\sigma \phi^{-1})^{m-1}$ and the equation 8.77 reduces to

$$-\left(\sigma \frac{R'}{R} - \frac{\phi'}{\phi}\right) [\lambda - \left((\beta R)^\sigma \phi^{-1}\right)]^{-1} + \sigma \frac{R'}{R} [\lambda - \left((\beta R)^\sigma R^{-1}\right)]^{-1} + \frac{\phi'}{\phi} = 0. \tag{8.79}$$

We have now to check that there exists a solution in λ of the above equation, where λ satisfies the additional condition 8.78. Let us define the function $\Phi(\lambda)$ as

$$\Phi(\lambda) \equiv -(\sigma\frac{R'}{R}-\frac{\phi'}{\phi})(\lambda-\frac{(\beta R)^\sigma}{R})+\sigma\frac{R'}{R}(\lambda-\frac{(\beta R)^\sigma}{\phi})+\frac{\phi'}{\phi}(\lambda-\frac{(\beta R)^\sigma}{R})(\lambda-\frac{(\beta R)^\sigma}{\phi}) \tag{8.80}$$

Note that λ^* solves 8.79 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

$$\Phi(\frac{(\beta R)^\sigma}{R}) < 0; \Phi(1) > 0. \tag{8.81}$$

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:

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

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 8.81 fails. For this purpose it suffices to show that

$$\Phi'(\lambda) = \frac{\phi'}{\phi}\left(1 + 2\lambda - \frac{(\beta R)^\sigma}{R} - \frac{(\beta R)^\sigma}{\phi}\right) > 0 \tag{8.82}$$

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 ?? 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.

Second and third best tax plans exist

In this section we provide a proof of the fact that the objects we have been discussing actually exist:

Proposition *For any* $\sigma \geq 0$ *and any* $\rho \in [-1, +\infty]$, *if* $\beta^\sigma R(\tau)^{\sigma-1} < 1$ *for every* τ, *then:*
(1) a second best optimal tax exists;
(2) a third best (incentive compatible) optimal tax exists.

Proof. Let $S \equiv [0,1]^N$, where N is the set of natural numbers, be the set of tax rates, endowed with the product topology (*i.e.* the topology of point-wise convergence), a compact set in this topology. As we have seen, the solution of the second best problem is equivalent to the solution of:

$$\max_{\tau \in S} \left(\frac{\sigma}{\sigma - 1}\right) H(\tau_0, \tau_1, \ldots). \tag{8.83}$$

If we define the subset T of all the tax sequences in S such that for all $t \geq 1$ we have

$$\left(\frac{\sigma}{\sigma - 1}\right) H(\tau_t, \tau_{t+1}, \ldots) \geq \left(\frac{\sigma}{\sigma - 1}\right) H(\bar{\tau}, \bar{\tau}, \ldots) \tag{8.84}$$

then the solution of the third best problem is equivalent to the solution of

$$\max_{\tau \in T} \left(\frac{\sigma}{\sigma - 1}\right) H(\tau_0, \tau_1, \ldots). \tag{8.85}$$

We now claim that the function H is uppersemicontinuous. The case $\sigma \geq 1$ is simple and we consider that first. By assumption, for any tax sequence and any s, $\beta^\sigma R(\tau_s)^{\sigma-1} < 1$, so the series defining the function Y is uniformly convergent for any sequence $(\tau_1, \ldots)$, hence the function Y is continuous. We also have that X is a sum of positive terms, and $X \geq 1$. It follows that H is continuous.

When $\sigma \leq 1$, the two series may diverge. But consider for any $\tau \in S$ the functions from the integers to the extended real line R^*, defined by

$$\xi_\tau(t) \equiv \prod_{s=1}^{t} \beta^\sigma \frac{R(\tau)^\sigma}{\phi(\tau)}; \gamma_\tau(t) \equiv \prod_{s=1}^{t} \beta^\sigma R(\tau)^{\sigma-1} \tag{8.86}$$

Take now each of the factors $\frac{R(\tau)^\sigma}{\phi(\tau)}$ and $R(\tau)^{\sigma-1}$. We claim that they are continuous functions from $[0,1]$ to R^*. In fact, they are continuous in the interior, and the behavior at 0 and 1 is as follows:

$$\lim_{\tau \to 0 (\tau \to 1)} R(\tau)^{\sigma-1} = +\infty, \text{ if } \rho \geq 0 \tag{8.87}$$

$$\lim_{\tau \to 0 (\tau \to 1)} R(\tau)^{\sigma-1} = a^{\frac{1-\sigma}{\rho}} (+\infty), \text{ if } \rho < 0 \tag{8.88}$$

while

$$\lim_{\tau \to 0 (\tau \to 1)} \frac{R(\tau)^\sigma}{\phi(\tau)} = +\infty, \text{ if } \rho \geq \frac{1-\sigma}{\sigma} \tag{8.89}$$

$$\lim_{\tau \to 0 (\tau \to 1)} \frac{R(\tau)^\sigma}{\phi(\tau)} = 0(+\infty) \text{ if } 0 \leq \rho \leq \frac{1-\sigma}{\sigma} \tag{8.90}$$

$$\lim_{\tau \to 0 (\tau \to 1)} \frac{R(\tau)^\sigma}{\phi(\tau)} = a^{\frac{-\sigma}{1-\rho}}(0) \text{ if } \rho < 0 \tag{8.91}$$

For all but the region where $0 \leq \rho \leq \frac{1-\sigma}{\sigma}$, we can now conclude that as $\tau_n \to \tau$, then $\xi_{\tau_n}(t), \gamma_{\tau_n}(t)$ converge to $\xi_\tau(t), \gamma_\tau(t)$ in the extended real line. Now by Fatou's lemma (with the counting measure on the integers) the functions X and Y are lower semicontinuous. Since they are both larger than or equal to 1, their product, the h function, is lower semicontinuous. Hence the function H is uppersemicontinuous.

For the region $0 \leq \rho \leq \frac{1-\sigma}{\sigma}$ we have to consider separately the case in which for some finite time t_0 the limit τ has $\tau_{t_0} = 0$ (otherwise the previous argument applies). In this case every element in the sum defining Y with index larger than t_0 converges to $+\infty$, hence since X is bounded below by 1, the product XY converges to $+\infty$.

Now the existence of the optimal plan for the second best problem is obvious. For the third best, note that if the function H is uppersemicontinuous then the set T is closed, hence compact. Now our claim follows for this case as well.

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.* Forthcoming.

[4] Benhabib, J. and Velasco, A.; (1996). On the optimal and Best Sustainable Taxes in an Open Economy. *European Economic Review.* Forthcoming.

[5] Benhabib, J. and Rustichini, A. ; (1995). Optimal Taxation without Committment, C. V. Starr working paper 96-18, New York University.

[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. anfd 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.