Optimal Taxes Without Commitment*

Jess Benhabib
Department of Economics, New York University,
269 Mercer St., 7th Floor,
New York, NY 10003

Aldo Rustichini
CentER, Tilburg University,
Warandelaan 2, P.O. Box 90153,
5000 LE Tilburg, The Netherlands.

March 11, 1998

*We would like to thank Michele Boldrin, Patrick Bolton, Christophe Chamley, V. V. Chari, Ken Judd, Larry Christiano, Chris Phelan, Philippe Weil, seminar participants at Carlos III, Columbia University, Université Libre de Bruxelles, and participants at NBER Conference on Economic Fluctuations at the Hoover Institute on February 10, 1996 for valuable comments. We also thank the C. V. Starr Center for Applied Economics at New York University for technical assitance.

Proposed Running Head: Taxes Without Commitment

Send Proofs to: Professor Jess Benhabib, New York University, Department of Economics, 269 Mercer Street, 7th Floor, New York, NY 10003

Abstract

In the problem of optimal taxation in an economy with labor and capital, the optimal solution when the government can commit to a sequence of tax rates entails the tax on capital to tend to zero in the limit, with all the tax burden on labor. It is well known, however, that this solution is time inconsistent; so if the commitment power is not perfect, this second best tax plan will not be sustainable.

We model explicitly the trade-off between the cost of revising the tax plan, and the benefit of the revision. As a result, when commitment is not possible, both the limit tax rate and the steady state capital are different from their levels in the second best solution. Limit taxes on capital may be strictly positive; but it may also be the case that the only sustainable plan has subsidies to capital. The subsidies induce an overaccumulation of capital, which becomes a commitment device against revisions of the tax plan. *Journal of Economic Literature* Classification numbers: H21, C73.

Key Words: Optimal Taxes, Commitment, Incentive Compatibility.

1 Introduction

In a competitive economy where capital and labor are used to produce output, a benevolent government chooses a sequence of tax rates on these two factors, trying to maximize the lifetime utility of a representative agent. In doing so it is facing, in the standard models in the literature (see for instance [8], [1], or [2]), two constraints. First it has to collect in each period a fixed, exogenously given, amount. It may use to this purpose the revenue from bonds to smooth the collection of revenue over time: but the overall payments have to be financed through tax collection. Second, since the economy is competitive, it has to anticipate that the public will adjust their choices to the chosen tax plan, by optimizing their consumption plans taking prices and taxes as given. Since the taxes will change the marginal returns on the factors, taxes will have a distortive effect, which will induce, over and above the loss of income, an efficiency loss. The plan chosen under these two constraints is called the *second best* tax plan. This is the dynamic version of the general problem of optimal taxation in a competitive economy: for a review see Mirlees [16].

The problem in the setup we have just described has been studied at length, both from the theoretical and empirical point of view.[1] In particular one result stands out for its simplicity, sharpness, and robustness: in the limit, the optimal tax rate on capital income is zero, and all the tax burden is on labor. We will refer to this as the Chamley-Judd result (see the papers by these two authors quoted above). In fact the sequence of tax rates has a bang-bang feature: capital is typically taxed maximally in the first periods, and then taxes are shifted on labor. So in the first phase, when the capital stock is a given quantity, distortion is minimized by taxing the stock factor (capital); in the second phase, where the distortive effects on capital would be great, taxes provide incentive to accumulation, and shift entirely on labor.

It is also well known, however, that the second best tax plan is *time inconsistent*: the government will have incentive to change the announced plan, *in the interest of the representative agent himself*, at later periods.[2] The reason for this inconststency is clear: in the early periods it is optimal to announce low tax rates on capital in the future, in order to promote accumulation. When the future becomes present, however, and capital has been accumulated, it becomes convenient to do the opposite and tax capital (with no distortive effect) rather than to impose distortionary taxes on labor. So, unless the government has some commitment power to bind itself to implement, at any future date, the plan it has announced in the very first period, the plan will not be credible. Such a commitment power is however hardly acceptable in this extreme form. In a more reasonable formulation, the government is aware that a change in plans may have costs, for example from

[1] See for example [5] or [15]. Several recent papers ([18], [6], [7] or[9]) treat the optimal taxation problem in the context of endogenous growth.

[2] For an early discussion of time-inconsistency see Kydland and Prescott [12].

the point of view of its own credibility. This considerations and these costs are not, however, present in the models we have referred to above.

In this paper we take the point of view that the constraint of time consistency should be taken into account, and explicitly modelled. (For some related work see [3],[4], [14], [13], or [20]) This is accomplished by imposing a third constraint on the choice of the government: the tax plan has to satisfy the additional restriction that in any period the utility that the representative agent will have, from that[19] point onwards, according to the announced plan, must be at least as large as the utility he would derive if the government changes its plans, no matter how this is done. We refer to this constraint as an *incentive compatibility constraint*, and denote the tax plans that satisfy it as *consistent plans*. In particular, in some of our formulations below, we take the extreme case in which after a deviation the government has a complete loss of credibility: from that point on the public will assume that it cannot commit to any policy which is not simply the period by period maximization of output. This equilibrium following a revision of the tax plan determines endogenously a deviation value, which may in particular depend on the value of the capital and other assets at the moment of deviation.

Now consider again the problem of designing the optimal taxes, and focus in particular on the steady state in the equilibrium with the optimal taxes. As mentioned above, the Chamley-Judd result tells us that in the limit the tax rate on capital income is zero. Therefore, the time inconsistency of the solution is most extreme precisely at steady state, when the tax on capital is smallest and the incentive to revise the plan in order to relieve the economy of distortionary labor taxes, the largest. This seems to rule out the promise of any low tax on capital in the far future, because it is a non-credible promise. But can the only optimal solution be an increase in the tax rate on capital?

Further reflection will show this conclusion is not obviously true. An change in the tax rate on capital may for instance reduce the accumulation rate and the steady state level of capital; and this in turn might reduce the utility per period to the representative agent at the steady state. Therefore in the choice of the sustainable optimal tax rate in the far future, the planning government will have to consider how the level of steady state capital will adjust to the different tax plans, and how the value of the two different options, sticking to the announced plan or modifying it, change as a consequence. Clearly a more detailed analysis is necessary, and this will be found in this paper.

It may be a surprise then that the optimal tax policy may have capital subsidies in the limit, as we show in sections 4.1, 4.2, and via calibrated examples in section 4.2.[3] This is surprising because even if the public does not believe a promise of a zero capital tax in the future, it may nevertheless find the promise of a future capital subsidy credible: the reason for this is that, once implemented the steady state capital stock would be higher, and this would reduce the incentive of the government to deviate from its promise. In such a situation a positive incentive to

[3] Judd [10] discusses how subsidies to capital at a steady state can be optimal under comitment if there are imperfectly competitive markets distorting th economy.

accumulate induces a high level of capital; both the utility per period and the value from deviating increase, but the first grows with capital more than the second, until the plan becomes sustainable. The higher level of capital has created an *endogenous commitment device* that makes the revision of the tax, conditional on that level of capital, too costly compared to the value of continuing with the announced plan. The economy has then reached a level of equilibrium in over-accumulation.

We may now proceed with a detailed presentation.

2 The Optimal Taxation Problem

In this section we will set up the problem to study the best sustainable taxes in the presence of government bonds. The elements of the vector $(c_t, L_t, k_t, b_t, w_t, r_t)$ represent respectively consumption, labor, capital, bonds, the wage rate net of labor taxes and the return on assets net of taxes, all at time t.

The utility of the representative agent is given by:

$$\sum_{t=0}^{\infty} \beta^t \left(u(c_t) - v(L_t)\right).$$

The government's problem is to maximize the agent's utility, but it must set labor and capital tax rates so as to create enough revenue to finance a stream of expenditures equal to G in each period, which is exogenous and fixed. Taxes on labor are the difference between the marginal product labor and w; and capital taxes are the difference between the marginal product of capital and r.The production function is $f(k, L)$. We assume:

Assumption 1 *The function u is concave, increasing, differentiable; the function v is convex, increasing and differentiable; $\beta \in (0, 1)$. The function f is concave, homogeneous of degree one.*

The government can buy or sell bonds to the public with rate of return r, which is the same as the rate of return on capital. We impose no constraints on the quantity of government bonds, which can be positive or negative: in principle the government can be a net debtor or creditor. However we restrict the net return on capital in the initial period to be non-negative: $r_0 \geq 0$.[4] This restriction prevents the government from taxing k_0 at a high enough rate to induce the public to borrow from it an amount sufficient to finance all future government expenditures from interest payments alone. Such a scheme would avoid distortionary taxes, but would also require a large amount of negative bonds and unacceptable levels of taxation of the inelastic initial capital. It is standard in the optimal taxation literature to rule out such a solution to the problem by bounding from below either r_0 , or b_0 and b.

[4] Of course zero is chosen for simplicity. Another lower bound may also do.

The problem of the government therefore is to set tax rates for the future to maximize the utility of the representative agent, subject to generating sufficient revenue to cover its expenditures. The problem can be formulated in the manner of Chamley or Judd, where the government chooses the sequence of vectors $\{(c_t, L_t, k_t, b_t, w_t, r_t)_t\}$ to satisfy the first order conditions of the agent, provided these are sufficient to yield the agent's optimum, and to meet its own revenue constraints. As we discussed in the introduction, the solution of this problem is time inconstistent.

2.1 The Optimal Consistent Taxation

We therefore formulate the problem differently. We investigate the best taxes that the government can announce for the future, subject to the constraint that it will not want to deviate from its announcement. Here "best" is again used in the sense of maximizing the utility of the agent, subject to generating sufficient revenue to cover expenditures G.

To define this problem we must specify what the consequences of deviating from announced taxes will be. For example, one consequence can be the loss of reputation for the government, which leads the public to expect maximal capital taxes in every period in the future, with the result that savings are reduced. In each period then the government must compare the benefits of deviating, which allows the government to impose lower distortionary labor taxes in that period, with the costs of deviating that comes from lower future savings rates and lower discounted utility for the agents. Therefore we impose on government's problem an incentive-compatibility constraint: at each future period, continuing with the announced tax policy must yield to the agent a discounted utility that is at least as high as the discounted utility that the agent will obtain if the government deviates in any way from its announced policy.

We will start by specifying a general functional form for the value of deviation that depends on the values of bonds and capital at the beginning of the period. A specific example, where in response to a deviation by the government the agents expect maximal taxes on capital in future periods and therefore stop investing, will be analyzed in section 4.2 below.

To summarize, the constraints imposed on the government by the first order conditions of the agent are given, for each $t = 1, 2...$, by:

$$w_t u'(c_t) - v'(L_t) = 0 \tag{2.1}$$

$$u'(c_t) - \beta r_{t+1} u'(c_{t+1}) = 0. \tag{2.2}$$

The budget constraint of the agent is given by :

$$-k_{t+1} - b_{t+1} + r_t(k_t + b_t) + w_t L_t = c_t \tag{2.3}$$

and the economy-wide feasibility constraint given by:

$$k_{t+1} + c_t + G - f(k_t, L_t) = 0. \tag{2.4}$$

The additional equations expressing the government budget constraint are automatically satisfied if the previous equations are satisfied; so we ignore them.

Let the value of deviation at time t be given by $V^D(k_t, b_t)$: we assume that this function is differentiable in both variables. The incentive compatibility constraints are expressed by the inequalities:

$$\sum_{t=i}^{\infty} \beta^{t-i} \left(u(c_t) - v(L_t) \right) \geq V^D(k_i, b_i) \tag{2.5}$$

for every $i = 0, 1, \ldots$ The Lagrangian of the government's problem is therefore:

$$\begin{array}{c} \sum_{t=0}^{\infty} \beta^t [u(c_t) - v(L_t) + \lambda_t \left(u'(c_t) - \beta r_{t+1} u'(c_{t+1}) \right) \\ +\mu_t \left(w_t u'(c_t) - v'(L_t) \right) + \eta_t \left(k_{t+1} + c_t + G - f(k_t, L_t) \right) \\ +\xi_t \left(r_t(k_t + b_t) + w_t L_t + G - f(k_t, L_t) - b_{t+1} \right) + \kappa_t r_t] \\ + \sum_{i=1}^{\infty} \gamma_i \left(\sum_{t=i}^{\infty} \beta^{t-i} \left(u(c_t) - v(L_t) - V^D(k_i, b_i) \right) \right) \end{array} \tag{2.6}$$

with respect to $(c_t, L_t, k_t, b_t, w_t, r_t)_t$, for a given pair k_0, b_0.

The vector $\{(\lambda_t, \mu_t, \eta_{t,} \xi_t, \kappa_t)_t\}$ is the vector of multipliers associated with the first order conditions for the optimization problem of the agent, the feasibility constraints and budget constraints; the γ_i' elements are the multipliers associated with the incentive constraints.

The maximand in the problem (2.6) above is not concave, even when the incentive constraints are ignored. This lack of concavity is well-known to arise in the standard formulations of the optimal taxation literature under commitment. Therefore the characterization results that are obtained yield only necessary conditions for optimality. For example, the well-known result due to Chamley and Judd states that if an optimal solution converges to a steady state, then the capital taxes asymptotically go to zero: this result follows directly from the first order conditions. Our analysis of the best sustainable tax program given below will also primarily lead to results that are obtained from the necessary conditions.

We now turn to the first order conditions for the problem given by (2.6). The conditions given below must hold at each $t = 1, 2...$ The first order condition with respect to k_t for the government's problem is:

$$-\eta_t f_K(k_t, L_t) + \beta^{-1} \eta_{t-1} + \xi_t \left(r_t - f_K(k_t, L_t) \right) = \gamma_t \beta^{-t} V_k^D(k_t, b_t)$$

where $f_K(k_t, L_t) = \frac{\partial f(k_t, L_t)}{\partial k}$and $V_k^D(k_t, b_t) = \frac{V^D(k_t, b_t)}{\partial k}$. This condition can also be expressed as :

$$\left(\xi_t + \eta_t \right) \left(r_t - f_K(k_t, L_t) \right) = \gamma_t \beta^{-t} V_k^D(k_t, b_t) - \beta^{-1} \left(-\eta_t + \eta_{t-1} \right) \tag{2.7}$$

The first order condition with respect to c_t is:

$$u''(c_t)(\lambda_t - \lambda_{t-1} r_t + \mu_t w_t) + u'(c_t) + \eta_t = -\beta^{-t} \left(\gamma * \beta \right)_t u'(c_t) \tag{2.8}$$

where we define $(\gamma * \beta)_t = \sum_{i=1}^{t} \gamma_i \beta^{t-i}$. The first order condition with respect to b_t, r_t, w_t, L_t respectively are:

$$-\xi_{t-1} + \xi_t \beta r_t = \gamma_t \beta^{-t} V_b^D(k_t, b_t), \tag{2.9}$$

where $V_b^D(k_t, b_t) = \frac{V^D(k_t, b_t)}{\partial b}$,

$$-\lambda_{t-1} u'(c_t) + \xi_t (k_t + b_t) + \kappa_t = 0, \tag{2.10}$$

$$\mu_t u'(c_t) + \xi_t L_t = 0, \tag{2.11}$$

$$(\xi_t + \eta_t)(w_t - f_L(k_t, L_t)) - \mu_t v''(L_t) - (\eta_t + u'(c_t)) w_t = \beta^{-t} (\gamma * \beta)_t v'(L_t) \tag{2.12}$$

and where $f_L(k_t, L_t) = \frac{\partial f(k_t, L_t)}{\partial L}$. The above equations, together with the initial conditions and transversality conditions, define the system to be studied. To analyze the properties of this system we first turn to the steady state.

3 The Steady State

We first note that when incentive constraints are ignored, the Chamley-Judd result that the optimal capital taxes at a steady state are zero follows immediately from the first order condition on capital. In fact in this case the equation 2.7 above simplifies to:

$$(\xi_t + \eta_t)(r_t - f_K(k_t, L_t)) = -\beta^{-1} \left(-\eta_t + \eta_{t-1} \right). \tag{3.13}$$

But at steady state the multipliers ξ_t and η_t are constant (see the discussion after the equation (5.56) in the appendix) hence the above equation implies $r = f_K$, which is the Chamley-Judd result.[5]. The remaining steady state equations are:

$$u''(c)(\lambda_t - \lambda_{t-1} r + \mu_t w) + u'(c) + \eta_t = -\beta^{-t} (\gamma * \beta)_t u'(c) \tag{3.14}$$

$$-\xi_{t-1} + \xi_t \beta r = \gamma_t \beta^{-t} V_b^D(k, b) \tag{3.15}$$

$$-\lambda_{t-1} u'(c) + \xi_t (k + b) + \kappa_t = 0 \tag{3.16}$$

$$\mu_t u'(c) + \xi_t L = 0 \tag{3.17}$$

$$(\xi_t + \eta_t)(w - f_L(k, L)) - \mu_t v''(L) - (\eta_t + u'(c)) w_t = \beta^{-t} (\gamma * \beta)_t v'(L) \tag{3.18}$$

[5]For an interesting analysis of the dynamics of redistributive taxation around a steady state in a special case, see [11].

$$wu'(c) - v'(L) = 0 \tag{3.19}$$

$$\beta r = 1 \tag{3.20}$$

$$k + c + G - f(k, L) = 0 \tag{3.21}$$

$$(r - 1)(k + b) + wL = c \tag{3.22}$$

In addition, if the incentive constraint is binding we also have,

$$(1 - \beta)^{-1} (u (f(k, L) - G - k) - v(L)) - V^D(k, b) = 0 \tag{3.23}$$

In those cases in which the incentive constraint is not binding we have the Chamley-Judd condition stated above:

$$\beta^{-1} = f_K(k, L) \tag{3.24}$$

Now using equations (3.19), (3.21) and (3.22), we obtain:

$$f(k, L) - G - k = (r - 1)(k + b) + (u'(f(k, L) - G - k))^{-1} v'(L)L$$

which can be written, using equation (3.20) as:

$$G = f(k, L) - b\left(\beta^{-1} - 1\right) - \beta^{-1}k - (u'(f(k, L) + G - k))^{-1} v'(L)L \tag{3.25}$$

This equation can be solved for (k, L), if b is given, either in conjunction with (3.24) in the case without incentive constraints, or in conjunction with (3.23) when incentive constraints are taken into account and are binding. In both cases it is clear that all the steady state values will depend on b, and therefore indirectly on both b_0 and k_0.[6] To analyze this dependence explicitly, we must consider the

[6] The dependence of steady state values on initial conditions when government bonds are present is well-known, and is discussed in Chamley [1]. The initial value of government bonds make a difference because the amount of revenue that the government needs to raise with distortionary taxes in order to finance its expenditures depends on its future revenues, that is its revenue steam resulting from its initial bond position. Note of course that for positive revenues the goverment must be a net creditor, and must buy bonds initially, which is highly unrealistic. Therefore a lower (non-positive) limit on initial government bonds, or an upper bound on the first period tax rate, which ammounts to the same, are reasonable assumptions. As discussed in Chamley [1] and [2] however, if there is a binding limit on the first period tax rate, the government will not find it optimal to accumulate assets in susequent periods by buying bonds from the public to the point of financing all its expenditures by interest collections. The reason is that the taxes after the initial period are distortionary, and there is a trade-off between distorting labor and capital markets early on versus distorting them later on. As Chamley shows, generally labor taxes will continue to be positive at the steady state even in an economy with bonds, provided there is a binding limit on first period taxes.

first-order conditions that apply to the initial period, and solve for the steady state values in conjunction with them. In the appendix we show how we can solve for the steady state values in this manner, given initial conditions (k_0, b_0). We define:

Definition 1 *Let $x(b) = (r(b), w(b), c(b), k(b), L(b))$ be a solution of the equations (3.19), (3.20), (3.21), (3.22), and (3.23). We refer to $x(b)$ as a candidate incentive-constrained steady state.*

As mentioned above, the steady state value of b must be determined in conjunction with initial conditions (k_0, b_0), as shown in the appendix.

Equation (3.25) can be used to solve for L in terms of (b, k), and in an economy without bonds in terms of k alone. The solution for L however may not be unique, as the special cases discussed in the following sections demonstrate. In such situations it will be necessary to determine which of the solutions for L in equation (3.25) is the optimizing one. This can be done by checking which of the various solutions of L yields the highest steady state values, given by the first term in equation (3.23). The solution of equation (3.23) itself however may have multiple solutions in (k, b), representing the intersections of the value of defection with the value of continuation along the optimal solution. Therefore the various combinations of steady state values, evaluated at the appropriate $L(k, b)$, have to be compared in order to determine the optimal stationary solution. This will also determine whether, at the optimal steady state satisfying the incentive constraints, the returns to capital will be taxed or subsidized. The next section will illustrate the various possibilities in special cases.

Determining the optimal steady state from among the possible steady state solutions above can be achieved through direct comparisons. A more analytical approach is to solve for the associated Lagrange multipliers and check whether they satisfy the appropriate Kuhn-Tucker conditions. In particular, for our problem given above, the multipliers associated with the incentive compatibility constraints, that is the multipliers given by the sequence $\{\gamma\}_i^\infty$, must be positive and summable[7]. In the proposition 1 below, we provide a condition to assure this. Define:

$$1 + a \equiv \frac{E - F}{E - F - D} = \frac{1}{1 - \frac{D}{E-F}}$$

where

$$\sigma_c \equiv \frac{u''(c)c}{u'(c)}, \sigma_L \equiv \frac{v''(L)L}{v'(L)}, \tag{3.26}$$

$$\frac{D}{E-F} = \left(\left(\frac{\left(V_k^D - r\right) M}{(f_K - r) w u'(\sigma_L + \sigma_c)} \right) + \left(\frac{\sigma_c V_b^D (k + b)}{u' (\sigma_c + \sigma_L) c} \right) (1 + \sigma_L) \right)^{-1}, \tag{3.27}$$

[7]More precisely they must be contained in the dual space of ℓ^∞, which contains ℓ^1, the space of summable sequences.

$$M \equiv (1+\sigma_L)\, w - f_L(k_t, L_t)$$

and where the above expressions, $\left(\frac{D}{E-F}\right)$ and M, are evaluated at steady state values. The following proposition is proved in the appendix:

Proposition 1 *Suppose that* $x(b)$ *is a solution to the government's problem (2.6) from initial conditions* $(k(b), b)$, *that is, it is a steady state solution for the problem given by (2.6). Then* $1 + a \in (0, \beta^{-1})$, *where* a *is evaluated at* $x(b)$.

In the special cases discussed in the next section, we will check whether the condition given by lemma (1) holds in order to rule out stationary solutions to equation (3.23) that are not optimal stationary solutions for the problem given above in (2.6)

4 Incentive Compatible Steady States

We have now all the elements needed to characterize the set of incentive compatible steady states; in particular we will investigate whether capital will be taxed or subsidized at the optimal incentive constrained steady states. For simplicity, we will start by focussing on the case of an economy without bonds. We give a partial characterization of steady states in lemma 1 and in the proposition 2. We obtain a sharper characterization in Proposition 3 for the case of a linear utility function. We then use the necessary conditions implied by the Lagrange multipliers, given in lemma 1, to rule out some steady states and identify the optimal ones. These results are in the two propositions 4 and 5. Finally, we provide a specific a family of examples where the optimal solution to the government problem can have an incentive constrained steady state solution at which capital is subsidized. Table 1 at the end of this section shows, in a parametrized and calibrated family of examples, how capital taxes and subsidies can vary with the curvature of the production function and the labor supply elasticity. Table 2 presents another family of examples, with a strictly concave utility function, and identifies the steady state with a capital subsidy as the optimal one, for equilibrium trajectories originating close to it. In the following we assume a stronger version of assumption 2:

Assumption 2 $v(0) = 0$ *and* $\lim_{L\to\infty} v'(L) = \infty$.

The procedure to determine the steady states is the following. Take any value k of the capital stock, and consider it as a possible candidate to be a steady state value. That is, consider the steady state equations (*i.e.* equations (4.29), (4.30) and (4.31) below) at that value of k. The solution of these equations determines the value of consumption and labor, and therefore of the utility of the representative consumer, at that steady state; this determines a function $W(k)$, which is the value per period to the representative consumer in an economy which maintains k as the steady state value, at the competitive equilibrium where the tax on capital income is zero, and enough tax revenue is generated to pay for G.

Now let $V^D(k)$ be the value of deviation at k. A necessary condition for k to be an incentive compatible steady state value is that the total future utility, $\frac{W(k)}{(1-\beta)}$, is at least as large as $V^D(k)$. So the candidate optimal steady states for the constrained problem are given by the intersection of the two curves described by V^D and $\frac{W}{(1-\beta)}$.

The analysis of an economy with bonds will be a straightforward modification of the analysis of this section: the equation (4.30) that follows should be replaced, in the analysis of the steady state of an economy with bonds, by the equation [8]

$$k\beta^{-1} + b(\beta^{-1} - 1) + wL = f(k, L) - G \tag{4.28}$$

The steady state values of (L, c, w), for a fixed steady state value of the capital stock, are determined by the three equations:

$$k + c + G = f(k, L) \tag{4.29}$$

$$k\beta^{-1} + wL = f(k, L) - G \tag{4.30}$$

$$u'(c)w = v'(L) \tag{4.31}$$

This system of three equations determines the values of the three unknowns (L, c, w); this solution however is generally not unique. Let us discuss the set of solutions. First we reduce the system to the single equation in the variable L:

$$k\beta^{-1} + \frac{v'(L)L}{u'(f(k,L) - k - G)} - f(k, L) + G \equiv \Phi(k, L) = 0 \tag{4.32}$$

Note that $\Phi(k, 0) = k\beta^{-1} + G$ and $\lim_{L\to\infty} \Phi(k, L) = \infty$ for every k; so the equation (4.32) has a (possibly empty) compact set of solutions. We denote by $L_1(k)$ and $L_2(k)$ respectively the smallest and largest solution of (4.32). By their definition:

$$\Phi_L(k, L_1(k)) \leq 0 \leq \Phi_L(k, L_2(k)). \tag{4.33}$$

Let now k_g be the steady state value of the capital stock in the case of commitment; as we know, this is the value determined in the system (4.29), (4.30), (4.31) above, plus the condition of zero tax on capital. This steady state will be a steady state of the problem without commitment if and only if $\frac{W(k_g)}{(1-\beta)} \geq V^D(k_g)$. If not, the steady state of the constrained problem will be, if it exists, one of the intersections of the two curves; let us denote it by k^*. (See Figure 1 where we use $V(k) \equiv \frac{W(k)}{(1-\beta)}$; k^* can be k_1 or k_2.)

[8] See a Section 5.1 in the appendix, and also footnote 3.

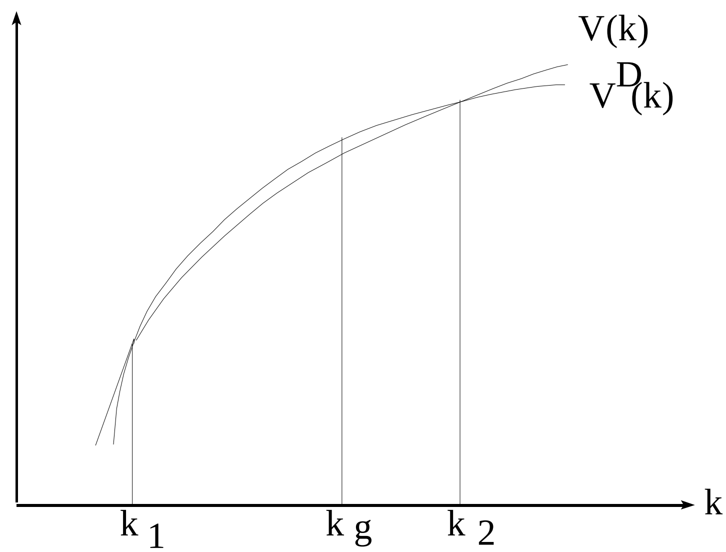

Figure 1: k_g is time-inconsistent

Figure 1 illustrates that perturbations of parameters of the system may change the intersections of $V(k)$ and $V^D(k)$, without necessarily changing the qualititative nature of the results. For example, if the value of deviation were to be higher because the trust of the public in the government to stick to its commitments could more easily and quickly restored after a deviation, the value of deviation would $V^D(k)$ would shift up. The two possible incentive compatible steady states would move apart, the lower one requiring a higher capital tax rate while the higher one requiring a higher capital subsidy. As long as the intersections of $V(k)$ and $V^D(k)$ are transversal, the two incentive compatible steady states which are distinct from the commitment steady state would be robust to perturbations. We further explore how the incentive compatible steady states change with various parameters in section 4.2 for some specific numerical examples (see Table 1 and 2).

Implicit differentiation of equation (4.32) gives:

Lemma 1 *In the region $k \geq k_g$, $L_2(k) \leq L_2(k_g)$; L_2 is a decreasing function; $f_K(k, L_2(k)) \leq \beta^{-1}$; and $f_K(\cdot, L_2(\cdot))$ is decreasing.*

Let us immediately note an implication of this:

Proposition 2 *Suppose that the constrained capital stock k^* is larger than k_g: then capital is subsidized at the steady state k^*.*

Of course we have to show that a situation as described in the proposition 2 may in fact occur. This is done the next section. We begin with a simple case of linear utilty, and further sharpen Proposition 2.

4.1 Linear Utility

In general the solution to the first order conditions for the problem of optimal time-consistent taxes may not be unique, and in particular may yield more than one equilibrium solution for the labor supply. Evaluating and comparing the multiple solutions to analytically identify the optimal one may be quite intractable. For this reason, we now restrict our attention to the case of utility functions linear in consumption, where $u(c) = c$. In fact a major simplification occurs in this case: the choice of the equilibrium quantity of labor is the largest solution of the first order conditions, both at steady state and at any point on the optimal path. The examples in the subsequent section involving subsidies to capital however will be constructed with linear as well as non-linear utility functions.

Take any pair (k_t, L_t), and set (this is imposed by the linear utility and the first order condition of the agent) $r_t = \beta^{-1}$; it is easy to see that there are still many possible equilibrium quantities for labor. However, if we let $\{(k_t, L_t, c_t, r_t, w_t), t = 0, 1, \ldots\}$ be the optimal path, then for any t, L_t is the largest solution of the equation in L given by:

$$k_t\beta^{-1} + v'(L)L + G - f(k_t, L) = 0 \tag{4.34}$$

In fact, if we let for any t the choice of k_t, k_{t+1}, r_t be given, the fact that the path is optimal implies that the value L_t is a solution of:

$$\max_L f(k_t, L) - v(L) - G - k_{t+1}$$

$$\text{subj. to } f(k_t, L) - v'(L)L - G - k_t\beta^{-1} = 0. \tag{4.35}$$

Now take $L_1 \leq L_2$ be any two solutions of the above constraint given by equation (4.35). The difference in the utility given by the two choices is equal to $v'(L_2)L_2 - v(L_2) - v'(L_1)L_1 + v(L_1)$. Our claim that L_t is the largest solution of the equation 4.34 follows from the fact that the function $v'(L)L - v(L)$ is strictly increasing.

Let $L_2(k_t)$ denote the largest solution of (4.35). It now follows that at the steady state the optimal choice of labor is $L_2(k)$ for any k. Also the lemma 1 has a stronger version, as follows.

Lemma 2 *The function L_2 is increasing in the region $k \leq k_g$, and decreasing in the region $k \geq k_g$; the function $f_K(\cdot, L_2(\cdot)) - \beta^{-1}$ is decreasing everywhere.*

The proof is straightforward, and based on the analysis of the system of differential equations giving L_2 and $f_K - \beta^{-1}$ as a function of k. Note that thanks to this lemma the steady state taxes on capital income are completely characterized:

Proposition 3 *If the steady state value k^* is larger than k_g , then capital is subsidized; if on the contrary k^* is smaller than k_g, then capital is taxed.*

We can now turn to the analysis of the necessary condition (4.36). As we have seen, a necessary condition for a steady state of the constrained problem, when the incentive compatibility constraints are binding, is that the multipliers computed in the previous section are positive and summable. In particular the condition that

$$1 + a \in (0, \beta^{-1}) \tag{4.36}$$

has to be satisfied. The following Lemma relates this condition to the slopes of $V^D(k)$ and $W(k)$.

Lemma 3 *1. If $V_k^D(k) \leq \frac{W'(k)}{1-\beta}$, then $1 + a \in (0, \beta^{-1})$ if and only if $V_k^D(k) \leq \beta^{-1}$.*
2. If $V_k^D(k) \geq \frac{W'(k)}{1-\beta}$, then $1 + a \in (0, \beta^{-1})$ if and only if $V_k^D(k) \geq \beta^{-1}$.

Proof. Recall that in the present case:

$$1 + a = \frac{1}{1 - R}; R \equiv \frac{v''(L)L(f_K(k, L) - \beta^{-1})}{(V_k^D - \beta^{-1})\Phi_L}$$

where $\Phi_L = \Phi_L(k, L_2(k))$; so that $1 + a \in (0, \beta^{-1})$ if and only if

$$\frac{v''(L)L(f_K(k, L) - \beta^{-1})}{(V_k^D - \beta^{-1})\Phi_L(1 - \beta)} \leq 1.$$

Now use the fact that

$$\frac{W'(k)}{1 - \beta} = \beta^{-1} + \frac{v''(L)L(f_K(k, L) - \beta^{-1})}{\Phi_L(1 - \beta)}$$

to derive our conclusion in both cases. ■

Proposition 4 *If the defection value function V^D is strictly concave, and*

$$\{k : V^D(k) \leq \frac{W(k)}{1 - \beta}\} = [k_1, k_2],$$

then it cannot be the case that $1 + a \in (0, \beta^{-1})$ both at k_1 and k_2; i.e. it cannot be that both are optimal steady states. (See Figure 1, where $V(k) \equiv \frac{W(k)}{(1-\beta)}$)

Note in fact that $V_k^D(k_2) \geq \frac{W'(k_2)}{1-\beta}$ and $V_k^D(k_1) \leq \frac{W'(k_1)}{1-\beta}$. From the lemma 3, if $1 + a \in (0, \beta^{-1})$ at k_2 then $V_k^D(k_2) \geq \beta^{-1}$, and therefore $V_k^D(k_1) > \beta^{-1}$. Now lemma 3 again implies that $1 + a \notin (0, \beta^{-1})$ at k_1. The argument in the other case is identical.

Lemma 4 *1. If $k^* \geq k_g$, and $V_k^D(k^*) \leq \frac{W'(k^*)}{1-\beta}$, then $1 + a \in (0, \beta^{-1})$.*
2. If $k^* \leq k_g$, and $V_k^D(k^*) \geq \frac{W'(k^*)}{1-\beta}$, then $1 + a \in (0, \beta^{-1})$.

Proof. Consider 1. By lemma 3 it suffices to show that

$$\frac{W'(k^*)}{1-\beta} \leq \beta^{-1},$$

and given the explicit form of $\frac{W'(k^*)}{1-\beta}$ it also suffices to show that

$$\frac{f_K(k,L)-\beta^{-1}}{\Phi_L} \leq 0.$$

But at $L_2(k)$ from lemma 3 we have that $f_K(k,L)-\beta^{-1} \leq 0$, and $\Phi_L(k, L_2(k) \geq 0$. The proof of 2 is similar. ■

We conclude that:

Proposition 5 *In the situation of proposition 4, if $k_g \geq k_2$, then the only constrained steady state which satisfies the necessary condition for optimality is k_2.*

4.2 Subsidies to Capital

In this section we analyze examples in which, at the constrained steady state satisfying first-order conditions, capital is subsidized at the optimal, constrained, tax plan. We use a specific value of deviation associated with irredeemable loss of reputation: once a government deviates, agents expect capital to be taxed at maximal rates at all future periods and therefore do not save. This is an extreme "punishment", and therefore the most likely to deter deviation. Milder forms, where the loss of reputation and confidence is temporary, or where expectations switch to higher rather than maximal tax rates, can also be analyzed along the lines of the example below.

We begin with a special example with linear utility, where:

$$u(c) = c, v(L) = \frac{1}{2}L^2; \tag{4.37}$$

and the production function is

$$f(k,L) = A(\epsilon)k + BL + \epsilon k^{\alpha} L^{1-\alpha}. \tag{4.38}$$

Later we consider a more general form, with non-linear utility, as in equation 4.39. Those examples should make clear that the capital subsidy results do not depend on the linearity of the utility function.

Here $A(\epsilon) = \frac{1}{\beta} - \varphi(\epsilon)$, where φ is a continuous strictly increasing function, with $\varphi(0) = 0$. The production function therefore is a weighted average of a component linear in capital and a standard Cobb-Douglas production function. This specification will be useful to construct parametric examples in Section 4.2.

It will be clear from the discussion that follows that the essential feature of the example is not the Cobb-Douglas form, but the fact that the marginal product of capital goes to infinity as the capital stock goes to zero.

The two functions $W(k)$ and $V^D(k)$ are defined as in the previous section. In particular the value of defection V^D has two components: the value from the first period, which depends on k, and the value from the following periods, which is a constant independent of k . We denote by $L^D(k)$ and $L^W(k)$ respectively the optimal labor choices for deviation and for the value of W. Note that $W(0) = V^D(0)$ and that $L^W(0) = L^D(0)$. The following lemma establishes that for $\epsilon > 0$, the slope of $W(k)$ becomes larger than the slope of $V^D(k)$ as $k \to 0$. Note also from the lemma that there is a discontinuity at $\epsilon = 0$.

Lemma 5 *1. As k tends to zero, the difference between the derivative of W term and the V^D term becomes, for $\beta > \alpha$, infinitely large and positive for any $\epsilon > 0$. More formally: for any fixed $\epsilon > 0$ and $\beta > \alpha$, and any number M, there is a value $k(\epsilon, M)$ such that for all values of k less than that, the difference is larger than M.*

2. For $\epsilon = 0$, the difference between the derivative of W and V^D tends, as k tends to zero, to a strictly negative value.

The proof of this lemma is relegated to the appendix. The lemma establishes that when $\epsilon > 0$ and $\beta > \alpha$ we have $\lim_{k\to 0}[((1-\beta)^{-1}W_k(k) - V_k^D(k)] > 0$, that is the slope of $(1-\beta)^{-1}W$ is strictly larger than the slope of V^D at $k = 0$. However if $\beta < 1$, then the function $(1-\beta)^{-1}W$ may again intersect the function V^D at positive values of k; so for values of k where $(1-\beta)^{-1}W(k) < V^D(k)$, k cannot be sustained as a steady state. A time-inconsistency problem arises if the optimal Chamley-Judd steady state for k_g falls within such a region; and the best incentive compatible steady state cannot lie in this region. Consider for example a situation as depicted in Figure 2 below, where $(1-\beta)^{-1} W(k) < V^D(k)$ for $k \in (k_1, k_2)$, and $k_g \in (k_1, k_2)$. The next proposition, that follows directly from lemma 4 establishes that both k_1 and k_2 are viable candidates for the optimal steady state.

Proposition 6 *For both $k = k_1$ and $k = k_2$, we have $1 + a \in (0, \beta^{-1})$.*

In the following section we present some calibrated examples to illustrate how the best sustainable steady states may have capital taxes or capital subsidies.

Some Numerical Examples

To get a better sense of the effects of incentive constraints on the tax rates and of how these effects vary with parameters, we resort to computations for a parametrized family of examples. We also modify the utility function (see equation 4.37) to allow the inverse labor supply elasticity, e, to become a parameter:

$$u(c) - v(L) = \frac{1}{1-\sigma}c^{(1-\sigma)} - \frac{1}{1+e}L^{1+e}; \tag{4.39}$$

We initially set $\sigma = 0$, so that the utility of consumption is linear. To define the production function we set $A(\epsilon) = \beta^{-1} - z\epsilon^r$, where $z = 1$, $r = 7/8$, $B = 3$ and

$\alpha = 0.34$. We set the discount factor at $\beta = 0.95$ and the government expenditures at $G = 2$. A higher ϵ or a higher z both increase the relative weight of the non-linear (Cobb-Douglas) part of the production function relative to its linear part.

In a number of cases it turns out there are 4 potential constrained steady states, that is intersections of $W(k)/(1-\beta)$ and $V^D(k)$, at $k = 0, k_1, k_2, k_3$. $W(k)/(1-\beta) - V^D(k)$ is positive on $(0, k_1)$, negative on (k_1, k_2), positive on (k_2, k_3) and negative again for $k > k_3$ (see Figure 2). In other cases, as the function $V(k)$ shifts down for certain parametrizations, the only constrained steady states are at k_1 and k_2. [9]

Table 1 and Figure 2 below illustrate how the constrained taxes differ from the taxes under commitment for various values of the inverse labor elasticity e, and for the parameter ϵ which is a measure of the curvature of the production function (see equation 4.38). In general capital is taxed at the steady state k_1 while it is subsidized at steady state k_2. The tax (subsidy) rates at candidate steady states k_1 and k_2 are quite close to 0, to some extent due to the linear utility of consumption and the almost linearity of the production function. As expected however these specifications also generate substantial differences in the steady state capital stocks at k_1, k_2 and k_g. We note that as the curvature of the production function is increased from $\epsilon = 0.00001$ to $\epsilon = 0.0025$ for a given value of e,the subsidy rate at k_2 increases while the tax rate at k_1 generally declines: it seems that when the incentive constraints bind and k_g cannot be sustained in the limit, more curvature in the production function favors lower taxes (or higher subsidies) to capital. In terms of Figure 2 this corresponds to a relative shift in both $V(k)$ and $V^D(k)$ so that both k_1 and k_2 increase. On the other hand as labor becomes more inelastic, that is as e is increased for a fixed ϵ, k_1 declines while k_2 increases (see Table 1). The same is true for the ratios $\frac{k_1}{L(k_1)}$ and $\frac{k_2}{L(k_2)}$ which determine the marginal products of capital: the former declines while the latter increases. In terms of Figure 2 this reflects a relative downward shift in the function $V(k)$. We also observe from Table 1 that as labor supply becomes more inelastic, the capital subsidies at k_2 increase: the labor taxes needed to finance a capital subsidy are less distortionary when the labor supply is more inelastic. However, since k_1 and $\frac{k_1}{L(k_1)}$ decline with e,while the marginal product of capital at k_g is always fixed at β^{-1}, capital taxes at k_1 must go up as labor supply becomes more inelastic.

Inspecting Table 1 below, two further observations emerge. First for $\{e = 0.5, \epsilon = 0.0025\}$ and for $\{e = 0.45, \epsilon = 0.002$ or $\epsilon = 0.0025\}$ we have $k_g < k_1$: this means that there would be capital subsidies at all of the positive candidate steady states k_1, k_2 and k_3. However in these cases it is clear that $V(k) > V^D(k)$ at k_g, so that k_g is unconstrained and can be sustained. Second, in Figure 2, the candidate steady state k_3 is apparent. But for $\epsilon = 0.0025$, when e goes from 1 to 1.15, the difference $W(k)/(1-\beta) - V(k)^D$ shifts down, so that k_2 and k_3 get close, merge and disappear, as also indicated in the Table 1. Then Lemma 3 implies that the steady state solution must be k_1, even though $k_g > k_1$. In fact lemma 3 indicates

[9] Note that unless $A(\epsilon)$ is larger than $1-\beta, V(k)$ eventually slopes down as the marginal product of capital becomes small and a high value of k becomes harder to sustain.

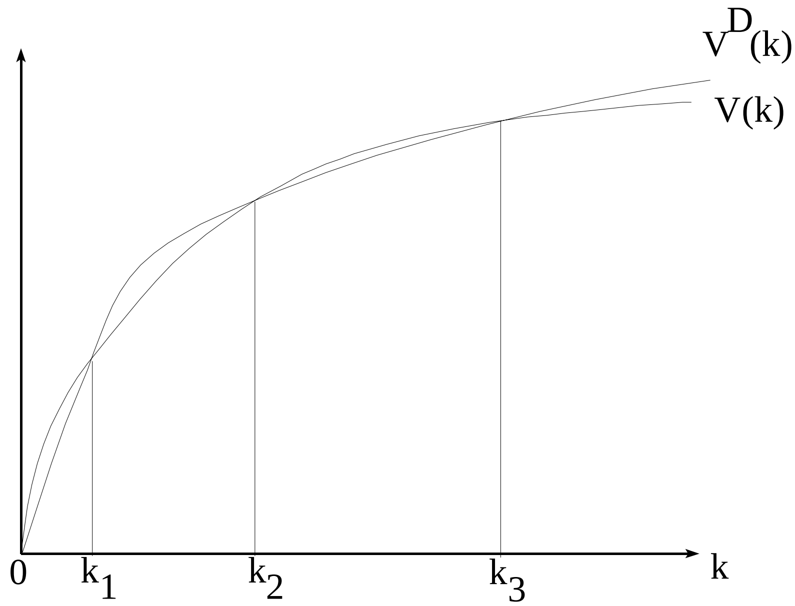

Figure 2: Multiple Steady States

that when $k_1 < k_g < k_2$, then either k_1 or k_2 satisfies the necessary conditions, but in this case there is no k_2.

e	ϵ	k_g	k_1	k_2	$L(k_g)$	$L(k_1)$	$L(k_2)$	$TAX(k_1)$	$TAX(k_2)$
1.15	.00001	.0001	.00001	6.05	1.6678	1.6678	1.6676	.00991	-.00005
1.15	.001	.0087	.0087	10.1	1.6683	1.6681	1.6428	.00839	-.00215
1.15	.002	.1000	.0236	24.4	1.6688	1.6686	1.5228	.00669	-.00401
1.15	.0025	.1043	.0324	—	1.6689	1.6689	—	.00592	—
1.00	.00001	.0046	.0002	5.7	2.0000	2.0000	1.9995	.00817	-.00004
1.00	.001	.1053	.0128	8.5	2.0005	2.0003	1.9827	.00704	-.00212
1.00	.002	.1199	.0340	13.2	2.0011	2.0009	1.9471	.00549	-.00393
1.00	.0025	.1250	.0464	17.7	2.0013	2.0012	1.9077	.00474	-.00482
.500	.00001	.1665	.00024	4.6	7.4641	7.4641	7.4640	.00317	-.00003
.500	.001	.3933	.1736	5.3	7.4658	7.4657	7.4586	.00164	-.00186
.500	.002	.4477	.4428	6.3	7.4677	7.4677	7.4519	.00003	-.00342
.500	.0025	.4668	.5933	6.9	7.4687	7.4686	7.4474	-.00074	-.00417
.450	.00001	.2129	.00039	4.3	9.8283	9.8283	9.8382	.00271	-.00003
.450	.001	.5179	.3028	5.0	9.8307	9.8306	9.8245	.00095	-.00175
.450	.002	.5895	.7902	5.6	9.8333	9.8332	9.8207	-.0006	-.00321
.450	.0025	.6147	1.073	5.9	9.8346	9.8342	9.8181	-.0015	-.00390

TABLE 1

In all the cases discussed above where $k_g \in (k_1, k_2)$, both k_1 and k_2 are potential steady state candidates: the condition given by lemma 1, that $(1+a) \in (0, \beta^{-1})$, is satisfied at both of them. The question arises therefore as to which of the steady states, given some initial conditions, is the optimal one, and whether a steady state with a capital subsidy can ever be optimal. It is possible to construct examples to show that a subsidized steady state can indeed be optimal. Consider for example a case for which the unconstrained (Chamley-Judd) steady state k_g is below, but arbitrarily close to the constrained steady state k_2. Consider the optimal path under commitment starting close to k_g, say at k_2, and which converges to k_g : clearly it dominates other feasible paths. Yet its value will now be arbitrarily close to a path which stays at k_2 with a subsidy, and which satisfies the incentive constraints. This means that the constrained path at k_2 will also dominate other feasible paths converging to k_1. Below we construct an example illustrating this possibility. We set the parameters as follows: $\alpha = 0.34, \beta = 0.95, B = 3, G = 2, r = 7/8$, $\epsilon = 0.005$, $\sigma = 0.2$ and $e = 0.5$. We let the parameter z in the definition of $A(\epsilon)$ vary. Note that in this case the utility of consumption is strictly concave, since $\sigma > 0$. Table 2 illustrates how as we vary z the optimal steady state under commitment, k_g, approaches k_2 from below and crosses it.

z	k_g	k_1	k_2	$L(k_g)$	$L(k_1)$	$L(k_2)$	$TAX(k_1)$	$TAX(k_2)$
0.5000	0.497	.1573	16.28	2.433	2.436	2.1500	.00522	-.00419
0.1023	5.382	.2042	5.391	2.3793	2.4371	2.3792	.00729	-.000001
0.1000	5.565	.2046	5.366	2.3772	2.4317	2.3796	.00730	.000023

TABLE 2

As shown in Table 2, for $z = 0.1023$, k_g and k_2 are very close. As we decrease z towards 0.1, k_g crosses k_2. Therefore a feasible path which starts at k_2 and stays there with a subsidy generates a value that is arbitrarily close to the value of an optimal path under commitment. Note also that at $z = 0.1$, although there seems to be a positive tax at both k_1 and k_2, the commitment steady state k_g exceeds k_2, so that in fact k_g can be sustained because $V(k_g) > V^D(k_g)$.

We can also explore the response of the constrained steady states to changes in σ which measures the curvature of the utility of consumption. Increasing σ increases k_2 and decreases k_1, which is equivalent to a downward shift of the function $V(k)$ relative to $V^D(k)$, which results in an increase in taxes at k_1 and an increase in subsidies at k_2. Further increases in σ eventually result in the disappearance of k_2 as it merges with k_3, and k_1 remains as the only positive constrained steady state. An analogous situation obtains for increases in z.

4.3 Capital as Commitment Device

Let us summarize the analysis of this simple model presented above. The commitment steady state k_g may not be sustainable and fall in between two sustainable steady states: one steady state is higher than the commitment steady state k_g, with negative taxes (subsidies) on capital; the second is lower, with positive taxes on capital.

The level k_g cannot be maintained because the promise of long run zero tax on capital is not credible; and this in turn happens because the temptation to revert to positive taxes is too strong. On the other hand, the promise, at the high steady state, of long-run *negative* taxes is credible. This may sound paradoxical. The reason for this is simply that capital subsidies give the appropriate incentive to accumulate a higher level of capital. As capital increases beyond k_g both the value of W and the value of V^D increase; they do, however, at different rates, so that eventually $\frac{W}{(1-\beta)}$ is above V^D. Note that at this higher capital stock the marginal product of labor tends to increase, and furthermore labor supply is lower, due the positive wealth effect, than it would be at k_g under commitment. (See lemma 1.) The one time advantage of taxing the existing capital stock at a higher rate in order to avoid the distortionary labor taxes is no longer worthwhile. So subsidies to capital become credible in the long run because they create a level of capital high enough to work as an endogenous commitment device against defection.

5 Appendix

5.1 The initial period

The first order conditions for the government's problem in the initial period are given by:

$$u''(c_0)\{\lambda_0 + \mu_0 w_0\} + u'(c_0) + \eta_0 = 0 \tag{5.40}$$

$$-\xi_0 + \xi_1 \beta r_1 = \gamma_1 \beta^{-1} V_b^D(k_1, b_1) \tag{5.41}$$

$$\xi_0(k_0 + b_0) + \kappa_0 = 0 \tag{5.42}$$

$$\mu_0 u'(c_0) + \xi_0 L_0 = 0 \tag{5.43}$$

$$(\xi_0 + \eta_0)(w_0 - f_L(k_0, L_0)) - \mu_0 v''(L_0) - (\eta_0 + u'(c_0))\, w_0 = 0 \tag{5.44}$$

$$w_0 u'(c_0) - v'(L_0) = 0 \tag{5.45}$$

$$u'(c_0) - \beta r_1 u'(c_1) = 0 \tag{5.46}$$

$$k_1 + c_0 + G - f(k_0, L_0) = 0 \tag{5.47}$$

$$-k_1 - b_1 + r_0(k_0 + b_0) + w_0 L_0 = c_0 \tag{5.48}$$

Equation (5.48) simplifies to:

$$-b_1 + r_0(k_0 + b_0) + \left((u'(f(k_0, L_0) + G - k_1))^{-1} v'(L_0) \right) L_0 = f(k_0, L_0) - G \quad (5.49)$$

We begin with the case without incentive constraints. First we solve for the steady state values in conjunction with the conditions for the initial period: we can therefore put $\gamma_i's$ equal to zero. Note that given b,the values of k and L are determined from (3.23) and (3.24): all other steady state variables, including steady state consumption, can be determined uniquely from these. To determine b note from (5.46) that since $\beta r = 1$, we also have $c_0 = c_1 \equiv c$: there is complete consumption smoothing due to the presence of bonds that yielding a constant return r from the initial period onwards. Now consider the transition to the steady state in one step: that is let the variables attain their steady state value in period one. We must check of course that this is consistent with a smoothed consumption level that is positive. If we put r_0 to its minimum value, which may be zero or $r_{\min}$, then the equations (5.45), (5.47) and (5.49), together with steady state equations (3.21), (3.24) and (3.25), determine $\{w_0, L_0, b, k, L, c\}$, where the unsubscripted variables represent the steady state values. Uniqueness of course is not assured. Note that putting r_0 to its minimum requires κ_0 to be non-zero, which from (5.42) requires ξ_0 to be non-zero and also from (3.15) requires ξ_i to be non-zero for all i. If ξ_0 is zero however, it implies that r_0 is interior, and that enough revenues to finance all future government expenditures can be generated without taxing initial capital at the maximum rate. It also implies that future labor taxes must be zero.[10] In the case where government bonds are not allowed, $\xi_0 = 0$ does not imply that $\xi_t = 0$ because equation (3.15), the first order condition with respect to bonds, is no longer applicable. In this case, while tax revenues from capital are sufficient to meet G in the initial period, in future periods labor taxes can still be positive because the government cannot rely on interest from selling bonds to the public. To avoid distorting incentives for saving, the government must tax labor to meet its expenditures on G after the initial period.

Solving for steady state values when incentive constraints are binding exactly parallels the procedure above, except that Chamley-Judd condition given by equation (3.24) is replaced by the equation (3.23), which specifies that the incentive constraints binding.

[10] This can be seen, when constraints are not binding and $\gamma_i's$ are zero, from the following: In equation 5.53 we can substuitute for η_t from equation 5.52 and solve for ξ_t as:

$$\xi_t = \frac{-u'(c_t)(f_L(k_t, L_t) - w_t)}{-(f_L(k_t, L_t) - w_t) + (\sigma_L w_t + \sigma_c f_L(k_{t,}, L_t))}$$

Clearly $\xi_t = 0$ implies that $f_L(k_t, L_t) - w_t = 0$ or that labor taxes are zero.

5.2 The multipliers

We still have to show that the steady state levels that have been computed satisfy the all necessary conditions associated with the multipliers, and that the multipliers are of the correct signs. Where multiple steady state solutions are found, only some may satisfy all these necessary conditions. We now turn therefore to solving for the values of the Lagrange multipliers on the optimal trajectory. The analysis that follows will also provide a proof of lemma 1. We start without making steady state assumptions.

We begin by solving for μ_t and λ_{t-1} from equations (2.10) and (2.11) and we substitute into $\{\lambda_t - \lambda_{t-1}r_t + \mu_t w_t\}$. We obtain, also using equation (2.3):

$$(\lambda_t - \lambda_{t-1}r_t + \mu_t w_t) = \frac{\xi_{t+1}}{u'(c_{t+1})}(k_{t+1} + b_{t+1}) - \frac{\xi_t}{u'(c_t)}(r_t(k_t + b_t) + w_t L_t) \quad (5.50)$$

Using (2.9) and (2.2) above we have:

$$\frac{\xi_{t+1}}{u'(c_{t+1})} = \frac{\xi_t}{u'(c_t)} + \frac{\gamma_t \beta^{-t} V_b^D(k_t, b_t)}{u'(c_t)} = \frac{\xi_t}{u'(c_t)} + \frac{\gamma_t \beta^{-t} V_b^D(k_t, b_t)}{u'(c_{t+1})\beta r_{t+1}} \quad (5.51)$$

We now can write, using (5.50) and (5.51), the equation (2.8) as:

$$\sigma_c \xi_t + u'(c_t) + \eta_t + \frac{u''(c_t)\left(\gamma_t \beta^{-t} V_b^D(k_t, b_t)\right)(k_{t+1} + b_{t+1})}{u'(c_t)} = -\beta^{-t}(\gamma * \beta)_t u'(c_t) \quad (5.52)$$

Solving (5.52) for $u'(c_t) + \eta_t$ and using (2.1) and (2.11), we can write (2.12) as

$$\begin{aligned} & (\xi_t + \eta_t)(w_t - f_L(k_t, L_t)) + \frac{v''(L_t)L_t\xi_t}{u'(c_t)} + \sigma_c \xi_t w_t \\ & + \left(\frac{u''(c_t)\left(\gamma_t \beta^{-t} V_b^D(k_t, b_t)\right)(k_{t+1} + b_{t+1})}{u'(c_t)}\right) w_t + \beta^{-t}(\gamma * \beta)_t u'(c_t) w_t \\ = \; & \beta^{-t}(\gamma * \beta)_t v'(L_t) \end{aligned} \quad (5.53)$$

or:

$$\begin{aligned} & (\xi_t + \eta_t)(w_t - f_L(k_t, L_t)) \\ = \; & -\sigma_c w_t \left(\xi_t + \left(\gamma_t \beta^{-t} V_b^D(k_t, b_t)\right)(k_t + b_t)(c_t)^{-1}\right) - \xi_t \sigma_L w_t \end{aligned} \quad (5.54)$$

Now dividing (2.7) by (5.54) we get:

$$\frac{(r_t - f_K(k_t, L_t))}{(w_t - f_L(k_t, L_t))} = \frac{\gamma_t \beta^{-t} V_k^D(k_t, b_t) + \beta^{-1}\left(-\eta_t + \eta_{t-1}\right)}{-\xi_t(\sigma_c w_t + \sigma_L w_t) - \sigma_c w_t \left(\left(\gamma_t \beta^{-t} V_b^D(k_t, b_t)\right)(k_{t+1} + b_{t+1})(c_t)^{-1}\right)} \quad (5.55)$$

We can now simplify equation (5.55) and solve it for ξ_t at steady state values. For this purpose we must evaluate $-\eta_t + \eta_{t-1}$ in a steady state. We can solve equation (5.52) for η_t, take first differences and using equation (2.9), and the fact that $\beta r = 1$, we can obtain the following equation:

$$\begin{aligned}
&-\eta_t + \eta_{t-1} \\
&= \sigma_c \left(\xi_t - \xi_{t-1}\right) + \left(\frac{u''(c)c}{u'(c)}\right)\left(\frac{k+b}{c}\right) V_b^D \left(\gamma_{t-1}\beta^{-t+1} - \gamma_t\beta^{-t}\right) + u'(c)\gamma_t\beta^{-t} \\
&= \sigma_c \left(\gamma_t\beta^{-t}\right) V_b^D + \sigma_c \left(\frac{k+b}{c}\right) V_b^D \left(\gamma_{t-1}\beta^{-t+1} - \gamma_t\beta^{-t}\right) + u'(c)\gamma_t\beta^{-t} \\
&= \left(\gamma_t\beta^{-t}\right)\left(u'(c) + V_b^D \left(\sigma_c \left(1 + \left(\frac{k+b}{c}\right)\left(\frac{\gamma_{t-1}\beta}{\gamma_t} - 1\right)\right)\right)\right)
\end{aligned} \tag{5.56}$$

Using (5.56) we can express (5.55) as:

$$\frac{(r - f_K(k,L))}{(w - f_L(k,L))} = \frac{\gamma_t\beta^{-t}\left(V_k^D(k,b) - \beta^{-1}\left(\left(u'(c) + V_b^D\left(\sigma_c\left(1 + \left(\frac{k+b}{c}\right)\left(\frac{\gamma_{t-1}\beta}{\gamma_t} - 1\right)\right)\right)\right)\right)\right)}{-\xi_t\left(\sigma_c w + \sigma_L w\right) - \sigma_c w\left(\left(\gamma_t\beta^{-t}V_b^D\right)(k+b)(c)^{-1}\right)} \tag{5.57}$$

Now solving (5.57) for ξ_t we get:

$$\xi_t = \gamma_t\beta^{-t}E \tag{5.58}$$

$$E = \left(\frac{\left(V_k^D(k,b) - \beta^{-1}\left(\left(u'(c) + V_b^D\left(\sigma_c\left(1 + \left(\frac{k+b}{c}\right)\left(\frac{\gamma_{t-1}\beta}{\gamma_t} - 1\right)\right)\right)\right)\right)\right)(w_t - f_L(k,L))}{-w\left(\sigma_c + \sigma_L\right)\left(\beta^{-1} - f_K(k,L)\right)}\right) - \frac{\sigma_c}{(\sigma_c + \sigma_L)}\left(V_b^D\right)(k+b)(c_t)^{-1} \tag{5.59}$$

Note that the expression E above depends on $\left(\frac{\gamma_{t-1}\beta}{\gamma_t}\right)$, which we will take to be equal to a constant, say $(1+a)$. We will demonstrate this to be the case below.

We store equation (5.58). Substituting η_t from (5.52) into (5.54)

$$\begin{aligned}
&-\left(1 + \beta^{-t}\left(\gamma * \beta\right)_t\right) u'(c)\left(w - f_L(k,L)\right) \\
&+ \left(\xi_t - \xi_t\sigma_c - \left(\frac{u''(c)\left(\gamma_t\beta^{-t}V_b^D\right)(k+b)}{u'(c)}\right)\right)\left(w - f_L(k,L)\right) \\
&+ \xi_t\sigma_L w + \sigma_c\xi_t w + + \left(\frac{u''(c)\left(\gamma_t\beta^{-t}V_b^D\right)(k+b)}{u'(c)}\right) w = 0
\end{aligned} \tag{5.60}$$

or

$$\left(\xi_t - \left(1 + \beta^{-t}\left(\gamma * \beta\right)_t\right) u'(c)\right)(w - f_L(k,L)) \tag{5.61}$$
$$+ \left(\xi_t \sigma_c + \left(\frac{u''(c)\left(\gamma_t \beta^{-t} V_b^D\right)(k+b)}{u'(c)}\right)\right) f_L(k,L) + \xi_t \sigma_L w = 0$$

Solving for ξ_t we have:

$$\xi_t = \frac{\left(1 + \beta^{-t}\left(\gamma * \beta\right)_t\right) u'(c)\left(w - f_L(k,L)\right) + f_L(k,L)\sigma_c\left(\gamma_t \beta^{-t} V_b^D\right)(k+b)(c)^{-1}}{\left((1+\sigma_L)w - (1-\sigma_c) f_L(k,L)\right)} \tag{5.62}$$

Using (5.58) equation (5.62) becomes :

$$\xi_t = \left(1 + \beta^{-t}\left(\gamma * \beta\right)_t\right) D + \gamma_t \beta^{-t} F = \gamma_t \beta^{-t} E \tag{5.63}$$

where:

$$D = \frac{u'(c)\left(w - f_L(k,L)\right)}{\left((1+\sigma_L)w - (1-\sigma_c) f_L(k,L)\right)} \tag{5.64}$$

$$F = \frac{f_L(k,L)\sigma_c\left(V_b^D\right)(k+b)(c)^{-1}}{\left((1+\sigma_L)w - (1-\sigma_c) f_L(k,L)\right)} \tag{5.65}$$

and E is as defined previously. Now since

$$\left(1 + \beta^{-t}\left(\gamma * \beta\right)_t\right) = 1 + \gamma_1 \beta^{-1} + \gamma_2 \beta^{-2} + \dots \gamma_t \beta^{-t}$$

equation (5.63) becomes:

$$\left(1 + \gamma_1 \beta^{-1} + \gamma_2 \beta^{-2} + \dots \gamma_t \beta^{-t}\right) D + \gamma_t \beta^{-t} F = \gamma_t \beta^{-t} E \tag{5.66}$$

Note however that at a steady state we can now compute the multipliers $\gamma_t \beta^{-t}$:

$$\begin{aligned} \gamma_t \beta^{-t} &= \left(1 + \gamma_1 \beta^{-1} + \dots \gamma_{t-1} \beta^{-(t-1)}\right)\left(\frac{D}{E-D-F}\right) \\ &= \left(1 + \gamma_1 \beta^{-1} + \dots \gamma_{t-2} \beta^{-(t-2)}\right)\left(\frac{D}{E-D-F}\right) + \gamma_{t-1}\beta^{-(t-1)}\left(\frac{D}{E-D-F}\right) \\ &= \left(\frac{E-F}{E-D-F}\right)\gamma_{t-1}\beta^{-(t-1)} = \left(\frac{E-F}{E-D-F}\right)^t \gamma_0 \end{aligned} \tag{5.67}$$

where γ_0 can be taken as unity: $\gamma_0 = 1$. We can also check whether the expression $\left(\frac{\gamma_{t-1}\beta}{\gamma_t}\right)$, appearing in the definition of E which was used in the derivation of $\gamma_t \beta^{-t}$ above, is indeed a constant as we had assumed. Using equation (5.67) we can compute $\left(\frac{\gamma_{t-1}\beta}{\gamma_t}\right) = \left(\frac{E-D-F}{E-F}\right)$, which is indeed constant at a steady state.

At a steady state, therefore, it is clear that E will be constant. It follows that ξ_t will have the same sign as $\gamma_t\beta^{-t}$. It also follows that $\gamma_t\beta^{-t}$ must be of constant sign, which requires that$\left(\frac{E-F}{E-D-F}\right) \geq 0$. Furthermore we must have $\gamma_t\beta^{-t}$ bounded and converging to zero since the Lagrange multipliers must be in summable, which requires that $\left(\frac{E-F}{E-D-F}\right) < \beta^{-1}$. We now investigate conditions under which $0 < \left(\frac{E-F}{E-D-F}\right) < \beta^{-1}$ at the steady state. Now let $1 + a \equiv \frac{E-F}{E-F-D}$. Therefore for $\gamma_t\beta^{-t}$ above to be bounded and summable, we have shown that we must have $1 + a \in (0, \beta^{-1})$. This proves Proposition 1 used in the text.

In particular for the case of a utility function that is linear in consumption ($\sigma = 0$), on which the results of section 4 are based, the expression for $1 + a$ can be further simplified. In such cases $1 + a \equiv \frac{E-F}{E-F-D}$ can be written as :

$$1 + a = \frac{((1+\sigma_L)\, w - f_L(k_t, L_t))}{((1+\sigma_L)\, w - f_L(k_t, L_t)) + \sigma_L w \left(\beta^{-1} - f_K(k_t, L_t)\right)\left(V_k^D(k_t.b_t) - \beta^{-1}\right)^{-1}} \tag{5.68}$$

In section 4, where we focus on the case of utility linear in consumption, we use the above expression for $1 + a$ to characterize the optimal steady states.

Proof of Lemma 5:

Proof of 1. Denote L^D and L^W respectively the equilibrium value of labor for deviation and for the value W. L^D is determined as follows: from the government budget constraint, setting $r = 0$, combining the two derivatives to get back the production function, and using the fact that $w = L$, we get:

$$G = f(k, L) - L^2.$$

For L^W: from the government budget constraint, setting the taxes on capital to zero,

$$G = BL + \epsilon(1-\alpha)k^\alpha L^{(1-\alpha)} - L^2.$$

When $k = 0$ the two values of labor are the same, denoted by

$$L(0) = 1/2[B + (B^2 - 4G)^{(1/2)}]$$

From the implicit function theorem,

$$\frac{dL^D}{dk} = -\frac{A + \epsilon\alpha(L/k)^{(1-\alpha)}}{B + \epsilon(1-\alpha)(k/L)^\alpha - 2L}$$

and

$$\frac{dL^W}{dk} = -\frac{\epsilon\alpha(1-\alpha)(L/k)^{(1-\alpha)}}{B + \epsilon(1-\alpha)^2(k/L)^\alpha - 2L}.$$

The values for the two programs are:

$$V^D(k) = f(k, L^D) - G - 1/2(L^D(k))^2 + \frac{\beta}{1-\beta}[BL_2(0) - G - (1/2)L_2(0)^2]$$

and

$$(1-\beta)^{-1}W(k) = (1-\beta)^{-1}[f(k, L^W) - k - G - 1/2(L^W(k))^2].$$

So if we now compute the limit of the difference between the slopes the two quantities as k tends to zero, we get:

$$\lim_{k\to 0}(1-\beta)^{-1}W_k(k) - V_k^D(k) = \lim_{k\to 0}\frac{\epsilon\alpha}{1-\beta}(L/k)^{(1-\alpha)}[\beta + \frac{B-L(0)}{2L(0)-B}(\beta-\alpha)].$$

We have used the fact that the ratio $\frac{L^W}{L^D}$ tends to 1 as k tends to zero. Also as discussed later we have ignored the A term in the numerator of the expression for L^D: this is correct when ϵ is non-zero, because the other terms (which become infinite) dominate. When $\beta > \alpha$, given the solution of $L(0)$ given above, the quantity $\lim_{k\to 0}(W_k - V_k^D)(k)$ is clearly positive, and tends to infinity as k tends to zero. This concludes the proof of the first point.

Proof of 2. Set $\epsilon = 0$ in the above computations, and get that:

$$\lim_{k\to 0}((1-\beta)^{-1}W_k - V_k^D)(k) = -(B - L(0))(\frac{A}{2L(0)-B}).$$

The intuition is again clear: the term A which is added to the increase in labor supply in the V^D case is dominated when ϵ is non zero by the other gains, which are all multiplied by ϵ. When however ϵ is zero this is the only term, because the relative gain for consumption becomes zero in the limit (since the linear terms cancel on both sides), L_k^W becomes zero, while L^D stays positive and makes V^D better. This concludes the proof. ■

References

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

[2] C. Chamley, Efficient Taxation in a Stylized Model of Intertemporal General Equilibrium. *Int. Econ. Rev.* **26** (1985), 451-468.

[3] V. V. Chari and P. J. Kehoe, Sustainable Plans. *J. Polit. Econ.* **98** (1990),783-802.

[4] V. V. Chari, L. J. Christiano and M. Eichenbaum, Expectation Traps and Discretion, working paper, (1996).

[5] V. V. Chari, L. J. Christiano and P. Kehoe, Optimal Fiscal Policy in a Business Cycle Model. *J. Polit. Econ.* **102** (1994), 617-652.

[6] L. E. Jones, R. E. Manuelli and P. E. Rossi, Optimal Taxation in Models of Endogenous Growth. *J. Polit. Econ.* **101** (1993), 485-517.

[7] L. E. Jones, R. E. Manuelli and P. E. Rossi, On the Optimal Taxation of Capital Income, working paper, (1995), Northwestern University.

[8] K. L. Judd, Redistributive Taxation in a Simple Perfect Foresight Model. *J. Pub. Econ.* **28** (1985), 59-83.

[9] K. L. Judd, Optimal Taxation and Spending in General Competitive Growth Models, working paper, (1995) Stanford University.

[10] K. L. Judd, The Optimal Tax Rate is Negative, working paper, (1995) Stanford University.

[11] M. C. Kemp, N. V. Long. and K. Shimomura, Cyclical and Noncyclical Redistributive Taxation, *Int. Econ. Rev.* **34** (1993), 415-430.

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

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

[14] A. Marcet, and R. Marimon, Communication, Commitment and Growth, *J. Econ. Theory* **58** (1992), 214-249.

[15] A. Marcet, T. J. Sargent and J. Seppala,Optimal Taxation Without State Contingent Debt, working paper, (1996).

[16] J. A. Mirlees, *The Theory of Optimal Taxation* , in *Handbook of Mathematical Economics*, vol. III, edited by K. J. Arrow and Intriligator, M. D., Elsevier Pub., North Holland, (1986).

[17] F. P. Ramsey, A Mathematical Theory of Saving. *Econ. J.* **38** (1928), 543-559.

[18] G. M. Milesi-Ferreti and N. Roubini, Growth Effects of Income and Consumption Taxes: Positive and Normative Analysis, N.B.E.R. working paper 5317, Cambridge, (1995).

[19] D. Pierce and E. Stachetti, Time Consistent Taxation by a Government with Redistributive Goals, *J. Econ. Theory* **72** (1997)**,** 282-305.

[20] N. L. Stokey, Credible Public Policy, *J. Econ. Dyn. Cont.* ***15** (1991), 627-657.*