ECONOMETRICA
JOURNAL OF THE ECONOMETRIC SOCIETY

WILEY

OPTIMAL TAXATION OF CAPITAL INCOME IN GENERAL EQUILIBRIUM WITH INFINITE LIVES

By Christophe Chamley[1]

This paper analyzes the optimal tax on capital income in general equilibrium models of the second best. Agents have infinite lives and utility functions which are extensions from the Koopmans form. The population is heterogeneous. The important property of the models is the equality between the social and the private discount rates in the long run. I find that the optimal tax rate is zero in the long run. For a special case of additively separable utility functions, I then determine the tax rates along the dynamic path and conditions that are sufficient for the local stability of the steady state.

INTRODUCTION

THERE IS CONSIDERABLE DEBATE in the tax policy community over whether or not to tax the income from capital and at what rates. The argument for a zero tax often rests on the disincentive effects of such a tax on the intertemporal allocation of resources. However, the studies of second best taxation indicate that the optimal capital income tax rates are not zero, except under quite restrictive assumptions about individual preferences. The main result of this paper is that when individuals have infinite lives and a utility function of a fairly general form, the optimal tax rate on capital income does tend to zero in the long run.

Previous studies of second best taxation in dynamic general equilibrium found different results because they relied on life-cycle models with limited scope for intertemporal substitution. In these earlier models, the individual's horizon was finite; in most of them, individual lives were reduced to two periods.[2]

The main assumption of this paper is that current consumption levels have a negligible effect on preferences among consumption plans sufficiently far in the future. The lifespan of individuals is infinite. No assumption is made about the interest elasticity of the supply of savings; it can take any arbitrary finite value, or be infinite.[3] This framework has interesting applications for an economy where individuals have an arbitrary utility function over a finite life and an altruistic concern for the utility level of their offspring.

In the paper I first consider the case where consumption levels have no impact on preferences in future periods. The general class of these utility functions has been characterized by Koopmans [11, 12]. In Section 1, I establish the zero

---

[2] Pestieau [13] has shown that under some conditions, the standard Ramsey tax formulae (Diamond and Mirrlees [8]), apply in the steady state of a general equilibrium model with overlapping generations. For a discussion of some of these conditions, see Atkinson and Sandmo [3].

[3] Summers [14] has remarked that in life-cycle models with additively separable utility functions and constant pure rate of time preference, the long-run interest elasticity of supply for savings increases with the number of periods. This would imply that the interest tax has a large welfare cost, and that its rate in the second best is relatively low. His analysis is restricted to a comparison between steady states. The framework of this paper does not make these restrictive assumptions about the utility function, and the second best policy optimizes over the entire dynamic path.

taxation of capital income in the steady state for this class of utility functions: in the intertemporal program of second best taxation, the tax rate on capital income tends to zero as the economy tends to a steady state, or a balanced growth path.

This result is first established in a heuristic model of intertemporal general equilibrium. A unique good is produced with capital and labor; it can be consumed or used to increase the capital stock. There is no lump-sum taxation; taxes on income from capital and labor are linear; and the tax rates can be optimized in each period. Finally, all agents are identical. These assumptions are not essential for the derivation of the result. The main property of the model which is used in the proof is the equality between the private and the social discount rate in the long run.

Section 2 shows how this equality and the validity of the result hold when some of the assumptions used in Section 1 are relaxed. The Koopmans form of the utility function is extended to allow for an effect of present decisions on future preferences: The utility function is separable between sequences of consumption programs in k consecutive periods, where k is an arbitrary fixed number. In this analysis, capital taxation is also equal to zero in the steady state, given economies with many consumption goods and with many agents of heterogeneous endowments and preferences. This latter property is particularly important for the relevance of the second best assumption. At the end of this section, I also discuss the implications of the result for the capital income tax and the bequest tax in economies with intergenerational transfers.

An essential part of the argument in this paper is the existence of a stable steady state for the dynamic path of second best. Therefore, Section 3 analyzes this problem for additively separable utility functions. Two results emerge.

First, the steady state is locally stable when one of two independent, sufficient conditions is satisfied. The first of these is a restriction on the shape of the utility function. The second is that the efficiency cost of taxation is relatively small. The latter condition seems to be applicable for a wider class of models. It implies that the dynamic path of second best is close to the path of first best. When the steady state is stable in the first best, it is also stable in the second best by a continuity argument.

Second, I find that the second best level of the capital income tax is fully described along the economy's dynamic transition path. At first, it is as large as constraints permit. It then abruptly switches, in finite time, to zero. This property is specific to the utility function which is considered here.

In Section 4, the policy implications of the results are summarized, and further extensions for future research are suggested.

1. THE OPTIMAL CAPITAL INCOME TAX IN THE LONG RUN

1.1. *The Model*

The private sector is represented by a single individual who has an infinite life. His utility function has the form proposed by Koopmans and his associates

[11, 12] and is defined as follows:

(1) $J(_tX) = U(x_t, J(_{t+1}X))$, $t \geq 1$,

with the notation: x_t is a vector of consumption of produced goods and leisure at time t, $_tX$ is the program $(x_t, \ldots, x_{t+k}, \ldots)$.

The function U is called the utility aggregator. The most restrictive property which is implied by the above functional form may be the limited noncomplementarity between periods (Koopmans [11]). This axiom will be partially relaxed below. The other important property is the stationarity of the utility function. It seems less restrictive in the present context since the main emphasis is on steady state results.

Without loss of generality one can assume that there is a unique produced good, and that the vector x_t represents the levels of consumption of the produced good c_t, labor supply l_t, and government consumption g_t, respectively:

(2) $x_t = (c_t, l_t, g_t)$.

The unique good can be used for consumption or as capital input in the production technology which is represented by the neoclassical function with constant returns to scale:

(3) $y_t = f(k_t, l_t)$.

The levels of output and capital are equal to y_t and k_t, respectively.

The variation of the capital stock between periods is equal to the difference between output and consumption:

(4) $k_{t+1} = k_t + f(k_t, l_t) - c_t - g_t$.

The government finances expenditures with linear taxes on the incomes of capital and labor. An equivalent assumption is that the fiscal instruments are the factor prices net of taxes, of capital and labor services in each period, $\bar{r}_t$ and $\bar{w}_t$, respectively. There is no restriction on the variations of the tax rates between periods. However, for a reason which will be clear below, the net rate of return $\bar{r}_t$ is constrained, and assumed to be greater than some arbitrary value, M. Without loss of generality, M can be chosen to be equal to zero:

(5) $\bar{r}_t \geq 0$ for all t.

The government can also match any imbalance between expenditures and revenues by issuing a debt which is perfectly substitutable with capital for the private individual (there is no uncertainty). If a surplus is generated, the government redeems the debt or buys capital. The variation of the debt b_t is equal to

(6) $b_{t+1} = (1 + \bar{r}_t)b_t + \bar{r}_tk_t + \bar{w}_tl_t - f(k_t, l_t) + g_t$.

The level of b_t is negative if the government has a positive wealth and owns capital.

Balanced budget policies can be analyzed in the same framework, with the additional restriction $b_t = 0$, for all t. In the remaining part of this section, one does not need to specify whether the budget is balanced in every period, and this constraint does not affect the main result presented below.

1.2. *The Second Best*

In the first period, the government announces the program of tax rates (or net factor prices), and expenditures. The representative agent is endowed with perfect foresight and behaves competitively: although the factor returns are endogenous to the decisions of saving and labor supply, they are considered to be independent of the agent's actions. For a given path of net factor prices $(\bar{r}_t, \bar{w}_t)_{t \geq 1}$, the representative agent chooses a program which maximizes the utility function (1), under the budget constraint. This program satisfies the first order conditions about the intratemporal and intertemporal choices, respectively,

$$(7) \qquad \bar{w}_t \frac{\partial U(c_t, l_t, g_t, J_{t+1})}{\partial c_t} + \frac{\partial U}{\partial l_t}(c_t, l_t, g_t, J_{t+1}) = 0,$$

$$(8) \qquad \frac{\partial U}{\partial c_t}(c_t, l_t, g_t, J_{t+1})$$

$$= (1 + \bar{r}_{t+1}) \frac{\partial U}{\partial J_{t+1}}(c_t, l_t, g_t, J_{t+1}) \frac{\partial U}{\partial c_{t+1}}(c_{t+1}, l_{t+1}, g_{t+1}, J_{t+2}),$$

where J_t is equal to $J(_tX)$. The levels of J_t and J_{t+1} are related by the equation:

$$(9) \qquad J_t - U(c_t, l_t, g_t, J_{t+1}) = 0.$$

As in the standard atemporal efficient tax problem with one person, the objective of the government is to choose values of the fiscal instruments such that the representative individual's utility is maximized. Constraints are imposed by the private sector's optimizing behavior and by the budget of the government; i.e., the value of b_t discounted in period one with the net rate of return, tends to zero when t tends to infinity.

The second-best problem is the maximization of the utility function $U(_1X)$ subject to the above constraints (4)–(9). The endogenous variables are the paths of c_t, l_t, g_t, J_t, $\bar{r}_t$, $\bar{w}_t$, k_t, and b_t for $t \geq 1$. Although the government does not directly control the levels of c_t and l_t, the framework is of second-best because the government chooses a program that satisfies the optimization restrictions (7)–(9) of the private agent. These restrictions correspond to the first order conditions of the private agent's program. His budget constraint is not taken as explicit because it is redundant in the formulation of the second best problem. When the dynamic path converges to a steady state, the budget constraints of the economy and of the government are satisfied (because of (4) and (6)). Therefore, the budget constraint of the private agent is also satisfied.

It is worth recalling that even in the first-best, the dynamic path that maximizes a utility function satisfying Koopmans' axioms does not always converge to a steady state, as shown by Iwai [10]. In this section, the convergence of the second-best program to a steady state is postulated. A discussion of this property is presented in the next section.

THEOREM 1: *When the representative individual's utility function has the form (1), and the second-best dynamic path converges to a steady state, the tax rate on capital income is equal to zero in this steady state.*

PROOF: Represent by $\tilde{\lambda}_t$ and $\tilde{\mu}_t$ the Lagrange multipliers associated with the constraints (4) and (6) respectively. The Hamiltonian is equal to

$$(10) \qquad H = U(_1X) + \sum_{t \geqslant 1} \tilde{\lambda}_t(-k_{t+1} + k_t + f(k_t, l_t) - c_t - g_t)$$

$$+ \sum_{t \geqslant 1} \tilde{\mu}_t(-b_{t+1} + (1 + \bar{r}_t)b_t + \bar{r}_t k_t + \bar{w}_t l_t - f(k_t, l_t) + g_t) + A,$$

where the terms in A do not include k_t as arguments.

We introduce the current value multipliers which are represented without a tilde:

$$(11) \qquad \tilde{z}_t = \left\{ \prod_{k=1}^{t-1} \frac{\partial U}{\partial J}(x_k, J_{k+1}) \right\} z_t, \quad t \geqslant 2,$$

where z_t represents one of the multipliers.

Differentiating H in (10), the first order condition with respect to k_{t+1} implies that

$$(12) \qquad \lambda_t = \frac{\partial U}{\partial J}(x_t, J_{t+1})[\lambda_{t+1}(1 + r_{t+1}) + \mu_{t+1}(\bar{r}_{t+1} - r_{t+1})],$$

where r_t is equal to the marginal product of capital. The discount factor (private and social) is defined by $\partial U(x_t, J_{t+1})/\partial J$, and is represented by δ_t. In the steady state all endogenous variables in (12) are constant over time. Omitting the time subscript,

$$(13) \qquad \lambda = \delta(\lambda(1 + r) + \mu(\bar{r} - r)).$$

The intertemporal first-order condition (8) of the private agent (in the steady state) is:

$$(14) \qquad 1 = \delta(1 + \bar{r}),$$

with the same value for δ as in (13). By combination of (13) and (14),

$$(15) \qquad (\lambda - \mu)(r - \bar{r}) = 0.$$

The variable μ represents the marginal social value of the public debt. It is also equal to the marginal value of replacing lump sum taxation by distortionary taxation (the marginal excess burden). In the second best this variable is negative (Atkinson and Stern [2]). The variable λ represents the marginal social value of capital. Therefore the term $\lambda - \mu$ is positive, and $\bar{r}$ is equal to r. Q.E.D.

There are two important steps in the proof which should be highlighted. First is equation (12) which has a simple interpretation: a marginal increment of capital in period t increases the quantity of available goods at time $t+1$ by the amount $1 + \bar{r}_{t+1}$, which has a social marginal value λ_{t+1}. In addition there is an increase

of tax revenues equal to $r_{t+1} - \bar{r}_{t+1}$, which enables the government to reduce its debt or other taxes by the same amount. The marginal reduction of the excess burden is equal to $-\mu_{t+1}(r_{t+1} - \bar{r}_{t+1})$. The sum of these two effects in period $t+1$ is discounted back by the discount factor, and is equal to the marginal value λ_t, in period t.

Second, the main step in the proof is the equality between the social and private discount rates for the steady state, in (13) and (14). This equality is satisfied for a wider class of models than those considered in this section. Some of these extensions are examined next.

2. EXTENSIONS AND APPLICATIONS

An important restriction of the Koopmans utility functions is the separability between periods (implied by the axiom of limited noncomplementarity), which rules out the formation of habits: the marginal rate of substitution between consumption levels in the future is independent of past actions.

Consider now the utility function defined by the expression:

$$(16) \qquad J(_tX) = U(\tilde{x}_t, J(_{t+1}X)), \quad t \geq 1,$$

where $\tilde{x}_t$ is the vector $(x_t, \ldots, x_{t+k})$, and k is an arbitrary constant integer. The term $_tX$ is defined in (1), where x_t is replaced by $\tilde{x}_t$.

This functional form allows for an effect of consumption in any period on the marginal rates of substitution in the k following periods. The application of the argument used in the previous section is straightforward. The result of a zero tax applies in this more general case.

Up to this point, the discussion has been conducted in a model of a one good economy, for the sake of simplicity. The extension to many consumption or capital goods is a simple exercise. If *ad valorem* taxes are available, their rates are in general different from zero under an optimal policy. Assume, for example, that there are two consumption goods, and that the function in (1) takes the form

$$(17) \qquad J_t = U(c_{t1}, c_{t2}, l_t, g_t, J_{t+1}),$$

where c_{t1} and c_{t2} represent the consumption levels of the two goods at time t. Assume also that the set of fiscal instruments is the wage tax, the tax on the rate of return, and an *ad valorem* tax on c_2. In the steady state the interest tax rate is nil, and the tax rate on c_2 is, in general, different from zero.

The functional form (17) is interesting because it can be interpreted as the utility function of a family where each individual lives two periods and has an operative bequest motive (Barro [4]). In this context, the values of c_{t1} and c_{t2} represent the levels of consumption in the two periods. The term J_{t+1} which represents the utility level of his next descendent, depends on the level of bequest. The combination of the tax on c_2 and of a general tax on the rate of return is equivalent to the combination of a tax on the rate of return for savings consumed during the life cycle and another tax on bequests. Theorem 1 implies that there should be no inheritance tax in the long run. Note that the result does not hold when the level of bequest is an explicit argument in the direct utility function

(as in Atkinson [1]). The intragenerational or internal discount rate (to use a terminology of Diamond [9]), is in general, different from the intergenerational or external discount rate. If a specific tax can be implemented on the interest income of savings used for life-cycle consumption, its rate is in general different from zero.[4]

From the institutional point of view, it is reasonable to consider the case where no distinction can be made between the savings for life-cycle consumption, and those accumulated for bequest. Theorem 1 implies that the zero tax on bequest overrides the efficient tax or subsidy on savings for life-cycle consumption. If there is a unique tax rate on the incomes of savings for life-cycle consumption and bequest, this rate is equal to zero in the long run.

The extension of the result to an economy with a finite number of heterogeneous agents, N, is a straightforward exercise. It is assumed that the individuals' utility functions are compatible with the existence of a steady state. (This problem also occurs in the first-best.) The government may use some uniform lump-sum transfers; it uses linear taxes on labor and capital income to maximize a social welfare function which, without loss of generality, is additive in the utilities of all agents. For the formulation of this problem of second-best, it is convenient to replace the capital accumulation equation (4) by the agents' savings equations:

$$(4a) \qquad a_{t+1}^i = (1 + \bar{r}_t)a_t^i + \bar{w}_t l_t^i - c_t^i \qquad\qquad\qquad (i = 1, \ldots, N).$$

The term a_t^i represents the level of assets of the agent i. It is equal to the sum of the holdings in capital and public debt.

The aggregate capital accumulation is then defined implicitly by the combination of these equations for private saving (in capital and public debt), and of the equation of the government deficit (6).

The same analytical method can also be applied when individuals have finite lives and are of two types. In the first, one's utility function is the sum of the utility over the life-cycle consumption and of the level of utility of the next descendant. Individuals of the second type do not leave a bequest and optimize over their own life cycle. When the social planner uses the same discount rate for the future life cyclers as the discount rate applied in the altruistic families, the long-run tax rate on capital income is equal to zero. This property does not depend on the initial distribution of human and nonhuman capital. However, it requires that individuals not be constrained at a corner solution for their bequest. (This may occur when the bequest cannot be negative.) The values of the tax rates on the transition path depend on the composition of the population.

One should emphasize that the zero interest tax result does not hinge on an infinitely elastic supply of savings. In the present framework, the discount rate is endogenous and depends on the levels of consumption and labor supply. In the steady state this rate is equal to the rate of return. Therefore, the long-run interest elasticity of the supply of capital of the private agent is, in general, not infinite.

---

[4] See Pestieau [13], or Atkinson and Sandmo [3].

Although the private and the social discount rates are always equal for a utility function (1), the zero tax result is generally valid only when the economy is in a steady state. It turns out, however, that for a specific class of utility functions the optimal interest tax is also equal to zero at least for all periods in the sufficiently distant future. We now turn to this specific class. The more restrictive framework will also enable us to solve the problem of the existence of a stable steady state.

3. THE DYNAMIC PATH—AN EXAMPLE

In this section, the complete dynamic path is analyzed for a specific class of utility functions. This context is somewhat restrictive. However, it generates a remarkable program for the interest tax. Also it facilitates the proof of the existence of a dynamic path which converges in the long run to a steady state.

For reasons of algebraic simplicity, time is continuous. The utility function of the private agent is assumed to be of the form:

$$(18) \qquad U = \int_0^\infty e^{-\rho t} \left[\frac{1}{1-\sigma} c_t^{1-\sigma} + L(1 - l_t) \right] dt,$$

where ρ and σ are fixed strictly positive parameters. When σ is equal to one, the term $c_{/(1-\sigma)}^{1-\sigma}$ is replaced by $\log c$. The function L is concave, and $L'(0)$ is infinite. Without loss of generality, the population is assumed to be constant.

For simplicity, the level of government expenditures is considered to be given and is not an argument in the utility function. The expenditures program $(g_t)_{t \geq 0}$ is assumed to converge to a constant level when time goes to infinity. (The convergence of the optimal program of expenditures could be proven by the method described below.) The fiscal instruments are the tax rates on the rate of return and the wage rate. In addition, the government can issue debt that is perfectly substitutable with capital (or it can save capital). The level of public debt at time zero is taken as given. A jump in the level of public debt at time zero would be equivalent to a lump sum transfer between the private agent and the government, and is ruled out in this second-best framework.

As in the previous section, an additional restriction is imposed on the admissible fiscal policies: the rate of return on assets, net of taxes, $\bar{r}_t$, is constrained to be nonnegative (another lower bound different from zero, could also be considered). The meaning of this constraint should now be clarified. At the beginning of time, the capital stock is in fixed supply and the constraint $\bar{r}_t \geq 0$ is obviously binding. Otherwise the efficient policy would be to implement a tax on capital such that the net rate of return is negative with an arbitrarily large absolute value during an infinitesimal interval of time. Asymptotically, this policy is equivalent to a lump sum taxation, and is obviously efficient.

In general one can expect from economic intuition that the constraint $\bar{r}_t \geq 0$ is binding during some interval of time starting at zero, until time t_1. After time t_1 the net rate of return should converge to the gross rate if the dynamic path converges to a steady state (according to Theorem 1). This description of the dynamic path applies for any utility function of the Koopmans type. The special

case of the additivity separable iso-elastic function (18) is remarkable. The transition from the regime of a high interest tax to that of a zero tax is instantaneous.

THEOREM 2: *Assume that the utility function of the private agent is of the form* (18) *(additively separable and iso-elastic in consumption). If the fiscal policy is efficient, there is a time τ such that for $t < \tau$, the constraint $\bar{r}_t \geq 0$ is binding, and for $t > \tau$, capital income is untaxed ($\bar{r}_t = r_t$).*

PROOF: Omitting the time subscript, the variations of the capital stock and the public debt are equal to

$$(19) \qquad \dot{k} = f(k, l) - c - g,$$

$$(20) \qquad \dot{b} = \bar{r}b + \bar{r}k + \bar{w}l - f(k, l) + g.$$

Represent by $u(c, l)$ the current utility function:

$$u(c, l) = \frac{1}{1 - \sigma} c^{1-\sigma} + L(1 - l).$$

Following a standard approach, the private individual's program imposes the following restrictions on the second-best policy:

$$(21) \qquad u'_c = q,$$

$$(22) \qquad u'_l = -q\bar{w},$$

$$(23) \qquad \dot{q} = q(\rho - \bar{r}).$$

The variable q represents the private marginal utility of assets. Using (21) and (22), c and l can be substituted in u as functions of $\bar{w}$ and q, to generate the function $u(c(\bar{w}, q), l(\bar{w}, q)) = v(\bar{w}, q)$.

The current value Hamiltonian of the second-best optimization problem is then equal to

$$(24) \qquad H = v(\bar{w}, q) + \xi q(\rho - \bar{r}) + \lambda(f(k, l) - c - g)$$
$$+ \mu(\bar{r}b + \bar{r}k + \bar{w}l - f(k, l) + g) + \nu\bar{r}.$$

The efficient policy satisfies the first order conditions:

$$(25) \qquad \dot{\xi} = \rho\xi - \frac{\partial H}{\partial q},$$

$$(26) \qquad \dot{\lambda} = \rho\lambda - \frac{\partial H}{\partial k},$$

$$(27) \qquad \dot{\mu} = \mu(\rho - \bar{r}),$$

$$(28) \qquad \frac{\partial H}{\partial \bar{w}} = 0,$$

$$(29) \qquad \frac{\partial H}{\partial \bar{r}} = -\xi q + \mu a + \nu = 0,$$

where a is the level of private assets, $a = k + b$; if $\nu > 0$, then $\bar{r} = 0$.

At time zero, there is no constraint on the level of q, and the associated costate variable ξ is equal to zero. One verifies that at time zero the value of ν is positive (it is equal to $-\mu a$, and the marginal tax excess burden $-\mu$ is positive). The constraint $\bar{r} \geqslant 0$ is binding, as expected.

Differentiating (29) and using (25), (23), (27), and the equation of private saving $\dot{a} = \bar{r}a + \bar{w}l - c$, the time variation of ν is equal to:

$$(30) \qquad \dot{\nu} = \rho \nu + G,$$

with

$$G = -qv'_q - \lambda(wql'_q - qc'_q) + \mu(w - \bar{w})ql'_q + \mu(c - \bar{w}l).$$

l'_q, c'_q, and v'_q represent partial derivatives with respect to q.

Because of the additive separability of u, the labor supply l depends only on the product $q\bar{w}$, and $\bar{w}l'_{\bar{w}} = ql'_q$. Using this relation and the condition $H'_{\bar{w}} = 0$, the expression of G can be rewritten:

$$(31) \qquad G = \frac{c}{\sigma}(q - \lambda + \sigma\mu).$$

Therefore the equation (30) is equivalent to

$$(32) \qquad \dot{\nu} = \rho \nu + \frac{c}{\sigma} Z, \qquad \text{with} \quad Z = q - \lambda + \sigma\mu.$$

From (23), (26), and (27), the time derivative of Z is equal to

$$(33) \qquad \dot{Z} = Z(\rho - \bar{r}) + (\lambda - \mu)(r - \bar{r}).$$

On an interval of time where the constraint $\bar{r} \geqslant 0$ is not binding, $\nu \equiv 0$. Therefore, the relation (32) implies that $Z \equiv 0$, and in (33) $r - \bar{r}$ is equal to zero (because $\lambda - \mu$ is positive).

The constraint $\bar{r} \geqslant 0$ cannot be binding forever (the marginal utility of private consumption q would grow to infinity since $\dot{q} = q(\rho - \bar{r})$, which is absurd). Call τ the greatest lower bound of t for which the constraint $\bar{r} \geqslant 0$ is not binding. To conclude the proof, it remains to show *ad absurdum* that $\bar{r} = r$, for $t > \tau$.

Assume that there is an interval (t_1, t_2) such that on this interval, $\bar{r} \geqslant 0$ is binding, and that t_1 is the lowest value of t greater than τ, for this to occur. At time t_1, $\nu = 0$ and $\dot{\nu} > 0$. In the relation (32), this implies that Z is strictly positive, which is impossible since Z is a continuous function of time and before time t_1, Z is equal to zero.   *Q.E.D.*

The previous result shows that there are two regimes for the interest tax. The policy is either to tax as much as possible or not at all. The tax has two effects. It raises revenues on existing capital but it also introduces intertemporal distortions in saving. The lump sum effect of the capital tax overrides the savings distortions for relatively small values of t ($t < \tau$). In the second regime, the saving

distortions become the predominant factor. The constraint $\bar{r} \geq 0$ is no longer binding. More important, there is no interest tax in the second-best, and the government generates revenues only by taxing wage income (or other commodities in a multi-good economy).

The duration of the first regime depends on the excess burden of these other taxes. Typically, when the value of the excess burden increases, the duration of the first regime τ, increases; the government extends the "tax recovery" period, accumulates wealth (or decreases the public debt), in order to lower the total discounted value of revenues generated in the second regime. (This is shown explicitly in an example considered below.)

The marginal value of the excess burden in the second-best is equal to the shadow price μ. It is simpler to express this value in units of private consumption, by the ratio μ/q. The comparison of the intertemporal first-order conditions of the private sector (23), and of the government (27), shows that they are identical, and that the marginal excess burden measured in units of private consumption is constant over time:

$$(34) \qquad \frac{\mu}{q} = \phi \qquad \text{is constant over time.}$$

This relation is simply an application of the principle according to which the government debt is used to smooth out variations of the excess burden between periods.[5]

The value of ϕ depends on the initial level of public debt. If the government owns initially a stock of capital equal to the present value of future expenditures, there is no need for (distortionary) taxation, and the marginal excess burden ϕ is equal to zero. In the standard case, the initial level of public wealth is smaller than the value of expenditures. Distortionary taxation is necessary, and the value of ϕ is negative (an increment of initial public debt generates a loss of social welfare). The relation between the values of marginal excess burden ϕ and of the public debt is made more explicit in the next results. For technical reasons, one has to make a distinction between the two regimes described in Theorem 2. Consider first the second regime which occurs after time τ.

THEOREM 3: *Let the marginal value of the excess burden ϕ, and of the capital stock at time τ, k_τ (where τ is defined in Theorem 2) be given, and at least one of the two (sufficient) conditions be satisfied: the absolute value of ϕ is small, or the third derivative of the function L is positive.*

Then, if the level of k_τ is sufficiently close to the steady state value of the capital stock, there is a unique value of the debt at time τ, b_τ, such that the dynamic path of second-best (if it exists), converges to the steady state, when t increases from τ to infinity.

[5] This principle has been analyzed by Barro [5] in the framework where the tax excess burden is, at each instant, an *ad hoc* quadratic loss function which depends on the revenues raised at the same instant.

The second sufficient condition is satisfied when the function L is isoelastic. The theorem is proven in the Appendix by determining the signs of the eigenvalues of the dynamic system, linearized at the steady state.

The first of the sufficient conditions is not surprising: when the marginal efficiency cost of taxation (measured by $|\phi|$), is small, the second-best solution is in the neighborhood of the first-best dynamic path, which is stable. Since the eigenvalues are continuous with respect to ϕ, the second-best steady state is also locally stable. This argument shows that the local stability of the second-best steady state must apply to a wider class of models than the one considered here.[6]

To complete the solution of the second-best problem, consider the first regime in the time interval $(0, \tau)$. Assume first that the values of ϕ, k_τ, and τ are given (and that Theorem 3 applies). There exists a unique set of values $(q_\tau, k_\tau, b_\tau, \xi_\tau, \lambda_\tau)$ such that the dynamic path converges to a steady state after time τ. The integration of the system of dynamic equations (23), (19), (20), (25), (26), back from time τ to time zero, determines the initial values $(q_0, k_0, b_0, \xi_0, \lambda_0)$, which depend on the given values of ϕ, k_τ, and τ. In this integration, the value of $\bar{w}$ is determined by (28), $\bar{r}$ is equal to its lower bound, zero by (29), and μ is equal to $q\phi$. The first three components of this vector are constrained: the values of k_0 and b_0 are exogenously given. The value of the shadow price ξ_0 is equal to zero because there is no constraint on the value of q at time zero (the equation (23) is only a constraint on the variation of q). These three constraints provide the additional equations which determine the value of ϕ, k_τ, and τ.[7] This completes the determination of the dynamic path.

It may be interesting at this point to have a rough estimate of the duration of the regime with interest taxation. Assume for tractability, that the gross rate of return r is equal to the discount rate ρ (near the steady state), and that the value of σ is equal to one. The value of τ is approximated by

$$(35) \qquad \tau = \frac{1}{\rho} \sqrt{-2\phi \frac{\alpha_0}{c_\tau}},$$

where a_0 is the level of private assets at time zero, and c_τ is the level of consumption at time τ (see Appendix, Section 2).

The value of τ is smaller (greater) than that of the formula (35) when the capital stock is relatively small (large) with respect to the steady state value. For example, take a value of $(\rho a_0/c_\tau)$ equal to $\frac{1}{4}$, a discount rate of four per cent per year and a marginal excess burden ϕ equal to $-\frac{1}{4}$. If the value of the gross rate of return r, in the interval of time $(0, \tau)$ is between ρ and 2ρ, the value of τ is

---

[6] See, for example, Chamley [7], and footnote 7. The saddle point property of the steady state in the second-best is not verified when ϕ is sufficiently large and L is negative.

[7] An argument which relies on the equality between the numbers of unknowns and equations may not convince the scrupulous reader. In the special case of fixed exogenous factor prices and ρ equal to one, there is a unique path converging to the steady state. The stability is global in this case. The proof is available from the author.

between 6 and 8.8 years. Of course this computation is only illustrative and the values of ϕ, c_τ, and τ should be determined simultaneously. It indicates, however, that the length of the period with capital income taxation at the 100 per cent rate can be significant.

4. CONCLUSION

Analysis of the capital income tax forces one to think about a number of difficult issues which are beyond the scope of a single study. The purpose of this paper is to consider only one aspect of the problem: how does the tax act as a fiscal tool in the general problem of second-best with perfect foresight and no uncertainty.

The results are simple and yet powerful. When the consumption decisions in a given period have only a negligible effect on the structure of preferences for periods in the distant future, then the second-best tax rate on capital income tends to zero in the long run.

This proposition does not contradict the standard properties of second-best taxation in the life-cycle framework, which depend on the structure of the life-cycle utility function and on a sufficient number of fiscal instruments. If there is a uniform tax on capital income and agents have an operational bequest motive, then the issue of long-run efficiency between generations prevails over the life-cycle efficiency problem. The (long-run) optimal interest tax is zero.

The capital income tax is also often regarded as a useful instrument for the redistribution of income. However, the analysis in this paper, applied to an economy of heterogeneous agents, shows that when individuals have long lives, a permanent tax on capital income is not an efficient policy for redistribution. This assumes that the social planner and the individuals use the same relative utility weights for intergenerational transfers.

The framework of this paper presents sufficient conditions for the zero interest tax. One property which is essential for the derivation of the result is the long-run equality between the social and the private discount rate. A useful extension of the present work would be to analyze the relation between these discount rates in other models of capital accumulation, and the implications of this relation for tax policy.

There is an obvious conflict between the efficiency of the capital income tax in the short run and its inefficiency in the long run. Under the assumptions of this paper, a commitment to no future tax on capital income is efficient. However, a government may be tempted to raise revenues by future levies on capital. Further studies will no doubt address this problem of time inconsistency and of the credibility of long-term policy commitments.

Harvard University

APPENDIX

1. PROOF OF THEOREM 3

The stability analysis is simplified when the level of labor l, is substituted for the net wage rate $\bar{w}$, as a control variable. The net wage rate is defined implicitly as a function of l and q, by the first order condition of the private sector:

$$(36) \qquad L'(1-l) = q\bar{w}.$$

By concavity of L, $\bar{w}'_l$ is positive. As shown in the text, when $\mu = \phi q$, ϕ is independent of time, and $\lambda = q + \sigma\mu$ (see the proof of Theorem 2, where $Z = 0$).

The Hamiltonian is now equal to

$$(37 \qquad H = u(c, l) + \xi q(\rho - \bar{r}) + \lambda(f(k, l) - c - g) + \mu(\bar{r}b + (\bar{w} - w)l + (\bar{r} - r)k - g),$$

where $\bar{r}$ is equal to r, c is a function of q, and $\bar{w}$ is a function of l and q (given in (37)). The first order conditions imply that:

$$(38) \qquad \begin{cases} \dot{k} = f(k, l) - c - g, \\ \dot{q} = q(\rho - r), \\ \dot{b} = rb + (\bar{w} - w)l - g, \\ \dfrac{\partial H}{\partial l} = 0, \end{cases}$$

$$(39) \qquad w - \bar{w} + \phi(\bar{w} - (1 - \sigma)w) + \phi l \bar{w}'_l = 0,$$

where $\bar{w}$ is a function of l and q (defined in (36)), and w is a function of k and l (by the production function).

This equation is of the form

$$(40) \qquad B(l, k, q) = 0,$$

and implicitly defines the value of l as a function of the state variables k and q, on the dynamic path. The partial derivatives of the function B are equal to:

$$(41) \qquad \begin{cases} B'_l = w'_l(1 - \phi(1 - \sigma)) + \bar{w}'_l(-1 + 2\phi) + \phi l \bar{w}''_{l^2}, \\ B'_k = w'_k(1 - \phi(1 - \sigma)), \\ B'_q = \bar{w}'_q(-1 + \phi) - \phi l \dfrac{\bar{w}'_l}{q}. \end{cases}$$

From the implicit definition of $\bar{w}$ in (36),

$$(42) \qquad \bar{w}'_q = -\frac{\bar{w}}{q}, \quad \bar{w}'_l = -\frac{L''}{q} > 0, \quad \text{and} \quad \bar{w}''_{l^2} = \frac{L'''}{q}.$$

The value of w'_l is negative and depends on the production function.

By existence of the second-best solution, the condition $B'_l < 0$ is assumed to be satisfied. (Note that it may be violated if $-\phi$ is large and $L''' < 0$ which implies $\bar{w}''_{l^2} < 0$, or when there is a relatively high degree of complementarity in both the production and the utility function, in which case the first term in B_l is positive and may be large.)

By linearization near the steady state (where the values of variables are denoted by an asterisk), the dynamic equations become

$$\begin{pmatrix} \dot{k} \\ \dot{q} \\ \dot{b} \end{pmatrix} = M \begin{pmatrix} k - k^* \\ q - q^* \\ b - b^* \end{pmatrix},$$

with the matrix M equal to

$$(44) \qquad M = \begin{pmatrix} r + wl'_k & wl'_q - c'_q & 0 \\ -qr'_k - qr'_l l'_k & -qr'_l l'_q & 0 \\ \cdot & \cdot & p \end{pmatrix},$$

where a dot represents a nonzero element. The partial derivatives of l with respect to k and q are defined implicitly in (40) and in the system

(41) $l'_k = -B'_k/B'_l$,   and $l'_q = -B'_q/B'_l$.

One eigenvalue of the matrix is equal to ρ and is positive. To prove that the two other eigenvalues have opposite signs, it is sufficient to show that the first determinant of order two Δ, is negative:

(45) $\Delta = -(r + wl_k)qr'_l l'_q + q(r'_k + r'_l l'_k)(wl'_q - c'_q)$.

Since $qc'_q = -c/\sigma$, $r'_l = -(k/l)r'_k$ (by the constant returns to scale assumption), and the level of output y is equal to $rk + wl$, after some manipulations,

(46) $\Delta = \dfrac{cr'_k}{\sigma}\left[1 + \dfrac{\sigma y}{c}\dfrac{ql'_q}{l} - \dfrac{kl'_k}{l} \right]$.

Substitute for l'_q and l'_k the ratios $-(B'_q/B'_l)$ and $-(B'_k/B'_l)$ of the expressions in (41) and use $w'_l = -(k/l)w'_k$, to find

(47) $\Delta = \dfrac{cr'_k}{\sigma}\dfrac{S}{B'_l}$,

with

$$S = \phi l\bar{w}'^2_l + \bar{w}'_l\left(-1 + \left(2 + \dfrac{\sigma y}{c} \right)\phi \right) - \dfrac{\sigma y}{c}\bar{w}'_q(-1 + \phi).$$

It follows immediately that either of the conditions given in Theorem 3 is sufficient for the determinant Δ to be negative. (Note also that if $-\phi$ is large and L''' is negative, the sign of Δ may be positive.)

The matrix M of the linearized system has two positive eigenvalues, and one negative eigenvalue. Therefore when the value of the constant variable ϕ is given, for each value of k near its steady level, there are unique values of q and b such that the dynamic system converges to the steady state.

The conditions of Theorem 3 are only sufficient conditions. The local stability could be proven under less restrictive assumptions. The previous analysis also shows that the existence of a unique second best dynamic path which converges to the steady state does not always occur.

2. THE LENGTH OF THE PERIOD OF CAPITAL INCOME TAXATION

The first order condition (26) implies that:

(48) $\dot{\lambda} = \rho\lambda - r\lambda - \mu(r - \bar{r})$.

Using (27), in the interval $(0, \tau)$,

(49) $\dot{\lambda} - \dot{\mu} = (\rho - r)(\lambda - \mu)$.

If r is equal to ρ, $\lambda - \mu$ is constant over time. According to (32), $\nu_t \equiv Z_t \equiv 0$ for $t \geq \tau$, and therefore $\lambda_\tau - \mu_\tau = q_\tau$, when σ is equal to one. This implies that

(50) $\lambda_t - \mu_t = \lambda_\tau - \mu_\tau = q_\tau$,   for $0 \leq t \leq \tau$.

In the interval $(0, \tau)$, $\bar{r} = 0$, the relation (23) reduces to $\dot{q} = \rho q$, and

(51) $q_t = q_\tau e^{\rho(t-\tau)}$.

Substituting $\lambda_t - \mu_t$ and q_t in (33), the expression (32) is equivalent to:

(52) $\dot{\nu}_t = \rho\nu_t + 1 - e^{\rho(\tau-t)}$,   for $0 \leq t \leq \tau$.

Since $\nu_0 = -\mu a_0$ in (29) ($\xi_0 = 0$), and $\nu_\tau = 0$, the value of τ is found by integration of (52), and is the solution of the equation

$$e^{\rho\tau} + e^{-\rho\tau} - 2 = -2\mu\rho a_0.$$

A good approximation is given by the formula (35) in the text.

When r is greater or smaller than some value θr, where θ is a constant, the equation in (49) can be replaced by an inequality. Upper or lower bounds can be obtained by a straightforward extension of the above method.

REFERENCES

[1] ATKINSON, A. B.: "Capital Taxes, the Redistribution of Wealth and Individual Savings," *Review of Economic Studies*, 38(1971), 209–227.

[2] ATKINSON, A. B., AND N. H. STERN: "Pigou, Taxation and Public Goods," *Review of Economic Studies*, 41(1974), 119–128.

[3] ATKINSON, A. B., AND A. SANDMO: "Welfare Implications of the Taxation of Savings," *Economic Journal*, 90(1980), 529–549.

[4] BARRO, R. J.: "Are Government Bonds Net Wealth?" *Journal of Political Economy*, 82(1974), 1095–1117.

[5] ———: "On the Determination of the Public Debt," *Journal of Political Economy*, 87(1979), 940–961.

[6] BRADFORD, D. F.: "Constraints on Government Investment Opportunities and the Choice of Discount Rates," *American Economic Review*, 65(1973), 887–899.

[7] CHAMLEY, C. P.: "Efficient Taxation in a Stylized Model of Intertemporal General Equilibrium," *International Economic Review*, 26(1985), 413–430.

[8] DIAMOND, P. A., AND J. MIRRLEES: "Optimal Taxation and Public Production: II—Tax Rules," *American Economic Review*, 61(1971), 261–268.

[9] DIAMOND, P. A.: "Taxation and Public Production in a Growing Setting," in *Models of Economic Growth*, ed. by J. A. Mirrlees and N. H. Stern. London: MacMillan, 1973.

[10] IWAI, K.: "Optimal Economic Growth and Stationary Ordinal Utility—A Fisherian Approach," *Journal of Economic Theory*, 5(1972), 121–151.

[11] KOOPMANS, T. C.: "Stationary Ordinal Utility and Impatience," *Econometrica*, 28(1960), 287–309.

[12] KOOPMANS, T. C., P. A. DIAMOND, AND R. W. WILLIAMSON: "Stationary Utility and Time Perspective," *Econometrica*, 32(1964), 82–100.

[13] PESTIEAU, P. M.: "Optimal Taxation and Discount Rate for Public Investment in a Growth Setting," *Journal of Public Economics*, 3(1974), 217–235.

[14] SUMMERS, L.: "Capital Taxation and Accumulation in a Life Cycle Growth Model," *American Economic Review*, 71(1981), 533–544.