

Published in Dahiya, S.B. (Ed.) (1999), *The Current State of Economic Science*, Spellbound Publications.

OPTIMAL DYNAMIC TAXATION

THOMAS I. RENSTRÖM

CentER for Economic Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. Email t.i.renstrom@kub.nl

This paper reviews the recent optimal dynamic tax literature, and links the results from dynastic one-person economies, dynastic heterogeneous individual economies and overlapping generations economies. The paper shows that the second best labour tax is positive, and further analyses the dynamic paths of capital and labour taxes, as well as the economy's adjustment under the optimal programme. Furthermore, we prove that in a heterogeneous individual economy, every individual's most preferred capital tax in steady state is zero. The optimal labour tax in a heterogeneous individual framework is similar to the optimal labour tax in a one-person economy.

JEL codes: H20, E62, D90

Keywords: Optimal taxation, dynamic taxation, fiscal policy

Acknowledgements: I am grateful to James Mirrlees for helpful comments.

1. Introduction

The concept of taxation in dynamic economies is not, at first sight, different from its static notion. We still distinguish between *first-best* and *second-best* analysis, and between *revenue-raising*, *redistributive*, and *corrective* taxation. At first we have two choices: either analysing an economy which in the absence of a government would be Pareto efficient (i.e. where the First Welfare Theorem holds), or an economy with imperfections (such as incomplete markets, imperfect competition, externalities). The role for taxation in the latter case is corrective. Of course we could combine them, however, but the issues should be explored in isolation to determine exactly which components are corrective. Whenever we analyse an economy which is perfect in absence of a government, we have two ways of introducing a government: for revenue raising or for redistribution. In the first case the government has to raise an exogenously given revenue, and is not allowed to use lump-sum taxes (usually referred to as Ramsey taxation). The second best tax system here would seek to arrange the taxes so as to minimise the distortions.

Ruling out lump-sum taxes here is ad hoc, since with one individual lump-sum taxation is optimal. Therefore the redistributive taxation framework with heterogeneous individuals is more attractive. Here the second-best arises because of an information asymmetry between the individuals and the government. The government cannot observe individuals' abilities or tastes, and therefore can only base the tax decisions on economic behaviour. Typically the government has access to lump-sum taxation as well (though not individual specific), and uses distortionary taxes for redistributive reasons (as in Mirrlees (1971)). A government funding requirement can of course be introduced as well.

There sometimes arises a direct conflict between the two ways of modelling second-best taxation. In the Ramsey-tax framework we would typically tax more those commodities which are relatively inelastic in demand.¹ In the framework with heterogeneous individuals, however, we should do the opposite if those commodities are consumed relatively more by poorer individuals (those with higher marginal utility of income).² Thus, ruling out lump-sum taxation on equity grounds and solving the optimal tax structure with a single individual, may give a tax structure that is not equitable at all!

Therefore, when we solve second-best tax problems in dynamic economies, the same conflict may appear, and we have to be careful in judging the results from one-person economies. However, as we will see in this paper, the principles obtained from second-best optimal-tax problems in one-person economies carry over to economies with heterogeneous individuals.

It is often possible to reinterpret a (static) commodity tax problem as a dynamic one. One could treat the different commodities as a single commodity consumed at different dates. Then, for example, a capital income tax turns out to be the same as taxing commodities at later dates at increasing rates. Therefore, the question of exempting savings from taxation would be a question of uniformity of consumption taxation (see Atkinson and Stiglitz (1972)).

However, there is a large difference between taxation in dynamic economies and static ones, related to time-inconsistency. The *second-best programme* would assume that the government can precommit to its future policy (i.e. that the government cannot in the future revise its original plan). The reason why a government would like to deviate from its

original programme in the future, is because the elasticities of the tax bases are different depending on when the policy decision is taken. For example, the elasticity of initial capital with respect to capital taxation today is zero. However, capital at a future date t is elastic with respect to a capital tax at time t if the decision on the tax is taken at an earlier date. But, once the government arrives at date t , capital is inelastic at t and the government would like to change the tax. Thus, if the government cannot make binding commitments to future tax policy, individuals will expect that the government is not going to follow the second-best plan. Individuals will base their expectations on what the government would find optimal to do at each date. This is the time-consistent equilibrium, and is usually referred to as the *third-best*, since it gives rise to a lower optimum.

We took the example of capital because dynamic taxation becomes most important in the field of optimal capital taxation. We can see how drastically different the policy prescriptions are. In a static framework, capital income is lump-sum income and should be taxed away. Thus taxing capital is desirable, or any tax reform which makes capital to bear the burden is desirable. In the dynamic-tax framework capital is foregone consumption, and individuals' incentives to save depend on the tax policy, and in fact it turns out that in most dynamic economies the optimal capital income tax in the long run is *zero* (e.g. Judd (1985), Chamley (1986), Renström (1998*b*)).

In a static framework the distinction second- and third-best policy never arises. The theory of dynamic taxation therefore offers a new direction of research: If governments cannot precommit to future policy what can be done to get closer to the second-best policy? Here we would analyse the institutional framework of tax policy. Governments solve their tax programmes without precommitment but under the constraint of constitutions. So here we have another difference: in a static framework constraining the planner only yields an inferior optimum, while in a dynamic economy such constraints may yield a higher optimum.

The purpose of this paper is to review the current literature on optimal dynamic taxation, and at the same time contribute with original results. We shall also suggest directions for further research.

We will begin by a brief literature review. Table 1 gives a taxonomy of research contributions on taxation in dynastic economies. They have some common characteristics:

- (i) The one-consumer economies are characterised by perfect competition and constant returns-to-scale in (aggregate) production, allowing for the First and Second Welfare Theorems to apply in absence of distortionary taxes. Thus, the focus is on how to "minimise" the distortions from the tax system, focusing on the efficiency aspects.³
- (ii) The individual (or family) has an infinite life, making it possible for the economy to reach a steady state. Also, in the differential game treatment by Kemp et.al., the infinite horizon allows for time independent feedback strategies.
- (iii) There is physical non-perishable capital, which is initially greater than zero. Some studies also allow for human capital, affecting labour productivity.
- (iv) The utility function is additively separable over time, so that consumption sufficiently distant in the past does not affect the marginal rates of substitution at the present.

Table 1 - Optimal Taxation in Dynastic Economies

	Consumers	Capital	Solution concept	Taxes
Judd (1985) (first part)	workers and capitalists	physical	full precommitment	capital
Chamley (1985)	identical	physical	full precommitment	labour
Chamley (1986)	identical	physical	full precommitment	capital, labour
Lucas (1990)	identical	physical, human	full precommitment	capital, labour
Kemp, van Long, and Shimomura (1993)	workers and capitalists	physical	full precommitment and no precommitment	capital
Correia (1996)	identical	physical,	full precommitment	capital, labour (fixed factor rents not taxable)
Jones, Manuelli, and Rossi (1997)	identical	physical, human	full precommitment	capital, labour consumption, labour equipment
Renström (1997)	identical	physical	full, partial, and no precommitment	capital, labour consumption

The studies by Judd (1985) and Kemp et. al. (1993) differ slightly from the others, in that two different classes of individuals are assumed. The focus is on *redistributive taxation*, rather than as a means of funding public goods.

The first part of Judd (1985) examines two cases (we postpone the discussion of the second part until later). Case I: workers supply labour inelastically and cannot borrow or lend. Capitalists own all capital and solve an intertemporal consumption-savings programme; the tax on capital income is redistributed only to the workers. Case II: both classes supply the same unit of work (inelastically) but differ in capital endowments. The capital income tax receipts are redistributed equal to both classes. Both classes are assumed to have the same rate of time preference, and in both cases the government has no funding requirement (of, say, public goods provision). Judd proves that the optimal capital-income tax (solving a Paretian welfare function) is zero in steady state. That this is true in Case I is perhaps most surprising. Judd gives the interpretation that the long-run tax on capital income represses wages, and therefore there is no gain from the redistributive capital income tax in the long run. Since the after-tax interest rate is equal to the rate of time preference at the steady state, labour bears all long-run burden of a capital tax. The optimal

tax programme à la Judd is clearly time inconsistent. If the government could reoptimise, it would not follow its original plan.

Not much research has been conducted on *time-consistent* fiscal policy.⁴ One important contribution is Kemp, Van Long and Shimomura (1993), who employ the basic assumptions of Judd's Case I, i.e. an economy with "workers" and "capitalists". They formulate the problem recursively, employing the methodology of *differential games*.⁵ Kemp et.al. use the *feedback Stackelberg* solution concept, with the government as (instantaneous) leader, and the capitalist takes the government's policy as given. They show that, indeed, the feedback equilibrium steady state is likely to involve positive taxation of capital income.

Chamley (1985) studies optimal wage-income taxation, exempting capital income from taxation, in a dynastic infinite-horizon economy. He concludes that the optimal wage tax is time inconsistent. The extension to tax also capital income is done in Chamley (1986).

Chamley (1986) may be divided into two parts. In the first part Chamley establishes the result that the optimal capital income tax is zero in steady state in an economy where the representative individual has preferences over private consumption, labour supply and public consumption, of the Koopmans (1960) form, which implies weak separability in consumption at different dates. The utility function is recursive and rules out habit formation (past consumption affecting current utility). An interpretation very often given is that under the assumptions made on individual preferences, uniform commodity taxation is optimal. That is, consumption at different dates should be taxed at the same rate, which translates into a zero capital-income tax. In the second part (which builds on Chamley (1985)) he derives the dynamic path for a particular utility function: additively separable in consumption and leisure and iso-elastic in consumption, assuming that the government expenditure path is exogenously given. Chamley reconfirms the zero capital-income tax result, and moreover shows that the economy would *initially* rely on capital income taxation at confiscatory rates (a maximum level of the tax rate is assumed to be 100%, otherwise the government would start confiscating property). After a period the policy switches to zero capital income taxation, i.e. the capital income tax reaches zero in *finite* time. However, this policy would be *time inconsistent* since, regardless of when the government solves the dynamic programme, it would prefer to set the capital income tax to the highest possible.

Lucas (1990) studied a similar economy to Chamley's. In addition he includes human capital and reconfirms the zero capital-income tax result at the steady state.

Jones, Manuelli and Rossi (1997) have a richer characterization of human capital accumulation. Individuals divide their time between market labour and learning. There are three goods in the economy. One consumption good, one good which increases labour productivity (e.g. gloves) and one good which increases learning ability (e.g. books), all purchased on the market. Taxes are levied on labour income, capital income, consumption expenditure and purchase of the labour-productivity increasing good. The good which increases learning ability remains untaxed. In some special cases the optimal long-run level is zero for all tax rates (i.e. a *zero-tax equilibrium*). The government builds up resources (capital) in the first periods, and rely on the returns from these resources for the entire future. Thus, in these cases, it is never optimal to smooth the excess burden over time. Since also Jones et. al. assume that the government can commit to all future taxes, their

optimal policy is time inconsistent.

Correia (1996) proved that when there is a factor (in addition to capital and labour) that cannot be taxed, then the optimal capital tax is generally not zero at the steady state.

To summarise: the second best optimal tax programme (when the tax structure is sufficiently rich) in one-person economies is generally characterised by a zero capital income tax in the steady state.

We shall now turn to a redistributive taxation framework with heterogeneous individuals.

Table 2 - Redistributive Taxation in Dynastic Economies

	Consumers	Decision	Solution method	Taxes
Judd (1985) (second part)	differ in preferences and capital endowments	optimal taxation	analytical	capital, labour, lump-sum transfer
Krusell, Quadrini, and Ríos-Rull (1996)	differ in skills and capital	voting on taxes, recursive formulation	numerical	capital, labour, consumption, income
Renström (1997)	differ in skills and capital	voting on representatives, full precommitment	analytical	capital labour consumption

The common characteristics of these studies are:

- (i) Individuals are heterogeneous and the focus is on the redistributive aspects of taxation. Therefore most studies abstract from government expenditure (except Judd (1985)) and assume that the tax receipts are redistributed lump sum to the individuals. Production (aggregate) are characterised by constant returns-to-scale.
- (ii) There is physical non-perishable capital.
- (iii) The utility function is additively separable over time, with exception for Judd (1985) where past consumption affect the rate of time preference [i.e. Uzawa (1968) preferences].
- (iv) The collective decision is one dimensional when the taxes are determined through majority voting.

The final part of Judd (1985) examines Pareto-efficient taxation in an economy with two individuals who differ in preferences. The preferences of individuals are of the Uzawa (1968) type, where the rate of time preferences are allowed to depend on past consumption and past labour supply. The government has access to a capital income tax, a labour income tax and an individual specific transfer (not depending on economic activity). The

government chooses the taxes for the entire future and the tax receipts are used for the individual specific transfers and for government expenditure. The government budget is period-by-period balanced. Because of the lump-sum transfer the economy is public-debt neutral. The government is assumed to maximise a weighted average of the two individuals' utilities. Judd proves that if the economy converges to a steady state the optimal capital income tax is zero if the shadow value of government expenditure is positive.

Krusell, Quadrini and Ríos-Rull (1996) analyse a dynastic economy with endogenous taxes by numerical methods. Individuals are of two types who differ in capital endowments and in productivity. Two types of experiments are conducted: (i) one tax rate and the lump-sum transfer are endogenous, (ii) two tax rates are endogenous and the lump-sum transfer exogenous. In both types of experiments there is period-by-period budget balance and a single-dimensional decision. They solve the dynamic politico-economic equilibrium by numerical methods and evaluate the equilibrium at the steady state. They conduct several experiments with different assumptions about the correlation between the skill distribution and the distribution of capital. The capital-income tax is not zero.

In Renström (1997) individuals have the same preferences but differ in labour productivity. The differences in productivity will generate different consumption and labour supply patterns and therefore different preferences over the tax rates. Individuals vote on government representatives (each individual being a candidate) and the majority winner chooses the policy for the entire future. It turns out that all individuals agree upon zero capital income tax in the steady state. The labour tax depends on the distance between mean and median skill.

Though both the above papers are median-voter models, we may view them as optimal tax models. The policy preferred by a median voter is the outcome of maximising a social welfare function where only the median person carry weight (a special case of social welfare function). Renström (1997) solves the second-best problem, and Krusell, Quadrini and Ríos-Rull (1996) the third-best. The latter contains results on the implemented taxes as a function of the constitutional setting (frequency of policy revisions, time lags in policy implementation). As mentioned earlier those restrictions may make the third best solution closer to the second best.

To summarise: the zero capital-income tax result is robust to the introduction of heterogeneous individuals.

Finally we shall give a brief overview of the optimal tax literature in overlapping generations economies, summarised in table 3.

They have some common characteristics (they are all applications of Diamond (1965)):

- (i) Production is characterised by perfect competition and constant returns-to-scale.⁶
- (ii) Consumers are typically heterogeneous, not only with respect to age.
- (iii) Consumers live for two periods and have no bequest motives.
- (iv) Physical capital is in the form of the old generations' savings.

Table 3 - Optimal Taxation in Overlapping Generations Economies

	Consumers	Popu- lation growth	Timing concept	Taxes
Diamond (1973)	differ in tastes	constant	recursive	linear on all commodities
Ordober and Phelps (1979)	differ in abilities	zero	recursive	non-linear on labour and capital
Atkinson and Sandmo (1980)	identical	constant	recursive	linear on labour capital,consumption
Park (1991)	differ in tastes, constant time endowments	constant	recursive	linear on labour, capital,consumption
Renström (1998b)	differ in skills and tastes	constant	open loop, recursive	linear on labour, capital,consumption

The optimality of zero capital income taxation at the steady state is obtained when no separability assumption in individuals' utilities is assumed, but with the social welfare function being weakly separable across generations (Koopmans (1960) form). This is only verified in Renström (1998b), which solves for the second-best (open loop) solution, while the other studies solve for the third-best (recursive) solution. We shall not review the overlapping generations economies further, the reader is referred to Renström (1998b). We should only notice that *the zero capital-income tax result carries over to overlapping generations economies.*

The paper is structured as follows. Section 2 describes and discusses the assumptions of the single-individual economy, and the individual economic behaviour is derived in section 3. Section 4 solves the second-best tax problem (perfect precommitment). The Chamley zero-capital-income tax is confirmed and interpreted, and it is proven that the labour tax is positive out of and in the steady state. The final part of section 4 is devoted to an analysis of the dynamic paths of the optimal taxes and the economy's transition dynamics under the optimal programme. Section 5 extends the analysis to heterogeneous individuals. Section 6 discusses the third-best tax programmes and proposes constitutional constraints that may make the third-best equilibrium closer to the second best. Section 7 concludes.

2. The Economy

2.1 Assumptions

A1 Individual's Preferences

The representative individual chooses consumption and labour supply paths, $c(t)$ and $l(t)$ for $t \in [0, \infty)$ so as to maximise her life-time utility

$$J(a_0) = \max_{c, l} \int_0^{\infty} e^{-\theta t} u(c(t), l(t)) dt \quad (1)$$

The instantaneous utility function u is assumed to be concave, and consumption and leisure are assumed to be normal goods.

A2 Individual's Constraint

The representative individual owns assets $a(t)$ (equal to the sum of outstanding public debt and the capital stock) and earns interest at a rate $r(t)$. For each unit of supplied labour he earns the wage rate $w(t)$. The taxes on capital income and labour income are denoted $\tau^k(t)$ and $\tau^l(t)$ respectively. Define the *after-tax returns* $\rho(t) \equiv [1 - \tau^k(t)]r(t)$ and $\omega(t) \equiv [1 - \tau^l(t)]w(t)$. Finally denote the consumption expenditure tax $\tau^c(t)$. The individual's budget constraint is therefore

$$\begin{aligned} \dot{a}(t) &= \rho(t)a(t) + \omega(t)l(t) - (1 + \tau^c(t))c(t) \\ a(0) &= a_0 \end{aligned} \quad (2)$$

A3 Production

There is a large number of competitive firms in the economy, each of whom operating under *constant-returns-to-scale* technology

$$f(k, l) = f_k(k, l)k + f_l(k, l)l \quad (3)$$

A4 Public Consumption

Real public expenditure takes the form of a sequence $g(t)$, $t \in [0, \infty)$, which is taken as exogenous.

A5 Government's Constraint

The government is assumed to be able to adjust fiscal policy in continuous time. The government is allowed to borrow and lend freely at the market rate of interest and takes the expenditure requirement $g(t)$, $t \in [0, \infty)$ as predetermined.

$$\begin{aligned} \dot{b}(t) &= \rho(t)b(t) - \tau^k(t)r(t)k(t) - \tau^l(t)w(t)l(t) - \tau^c(t)c(t) + g(t) \\ b(0) &= b_0 \end{aligned} \quad (4)$$

Finally the economy is assumed to be endowed with some initial capital $k_0 > 0$. The evolution of the capital stock is therefore

$$\begin{aligned}\dot{k}(t) &= f(k(t), l(t)) - c(t) - \mathbf{g}(t) \\ k(0) &= k_0\end{aligned}\tag{5}$$

2.2 On the Assumptions

A1. The utility function (1) is the discounted flow of instantaneous utilities. The constant discount rate implies that utility is additively separable in commodities at different dates. This makes marginal rate of substitution between two dated commodities independent of consumption at all other dates. Applying the analysis by Atkinson and Stiglitz (1972), this would imply, in the absence of individuals owning initial assets ($a_0=0$) that capital income should be exempted from taxation (uniform consumption taxation), if, and when, consumption is constant over time (as it is in steady state).

The infinite horizon assumption enables us to have a steady state where all quantities are constant for the representative consumer. This assumption is motivated by thinking of an economy where individuals have finite lives but have bequest motives such that the utility of an offspring enters the utility of the parent [Barro (1974)].

A2. The budget constraint (2) implies that capital markets are perfect: the individual can borrow and lend at the same market interest rate, without any constraints. The assumption is plausible in the type of economy we study. If we were to incorporate borrowing and lending constraints we would have to explicitly model the source of such constraints (asymmetric information or other market imperfections).

A3. Perfect competition, constant-returns-to-scale technology, and no production externalities. If there was imperfect competition, economies of scale, or production externalities, the tax rates would involve a "Pigouvian" element. The optimal taxation framework here is concerned with minimising the distortions caused by taxation, while the Pigouvian approach would, on the other hand, deal with corrective taxation.

A4. Most often in Ramsey-tax problems government expenditure (in real terms) is exogenously determined. We may motivate this assumption by thinking of the government's problem in two steps. First, for any *given* public expenditure decision the tax structure has to be *optimal* (the Ramsey problem). Second, given the optimal tax structure for each level of public expenditure, the government chooses its preferred spending taking into account the optimal tax structure. Thus, endogenising the public expenditure would not change the *tax rules*.

A5. There is tax equivalence in this economy. The degree of equivalence is such that *exactly one* of the tax paths $\tau^k(t)$, $\tau^l(t)$, and $\tau^c(t)$, ($0 \leq t < \infty$), may be normalised (to zero or to something else). This is a general property of dynamic economies, the degree of equivalence remain when the standard economy is extended to allow for other consumption goods as well. See Renström (1996a) for a formal analysis on these issues. We will normalise the consumption tax rate to zero, $\tau^c(t)=0 \forall t$, for most of the analysis, to keep the similarity with previous work.

3. Individual Economic Behaviour

In this section we will solve for the individual's economic behaviour, taking as given arbitrary tax paths. Setting $\tau^c(t)=0$ we may write the current-value Hamiltonian for the representative individual as

$$H = u(c(t), l(t)) + q(t)\{\rho(t)a(t) + \omega(t)l(t) - c(t)\} \quad (6)$$

The first-order conditions

$$u_c(c(t), l(t)) - q(t) = 0 \quad (7) \qquad u_l(c(t), l(t)) + q(t)\omega(t) = 0 \quad (8)$$

$$q(t)\rho(t) = \theta q(t) - \dot{q}(t) \quad (9)$$

describe the individual's choice of c and l as functions of the co-state q , up to the initial value of q , i.e. $q(0)$. The initial value of the co-state (i.e. marginal utility of the state [individual's assets]) is chosen to its lowest possible value subject to the intertemporal budget constraint

$$0 = a(0) + \int_0^{\infty} e^{-\int_0^t \rho(s) ds} [\omega(t)l(t) - c(t)] dt \quad (10)$$

Therefore, $q(0)$ depends on *all* future tax rates.

Equations (7) and (8) form a system such that c and l may be solved for as functions of q and ω . Their partial derivatives are obtained by differentiating through (8) and (9)

$$\partial c(t)/\partial q(t) = [u_{ll}(t) + \omega(t)u_{cl}(t)]D^{-1} \quad (11) \qquad \partial c(t)/\partial \omega(t) = u_{cl}(t)q(t)D^{-1} \quad (12)$$

$$\partial l(t)/\partial q(t) = -[u_{cc}(t)\omega(t) + u_{lc}(t)]D^{-1} \quad (13) \qquad \partial l(t)/\partial \omega(t) = -u_{cc}(t)q(t)D^{-1} \quad (14)$$

where $u_{ll}(t)$ is shorthand for $u_{ll}(c(t), l(t))$ etc, and $D \equiv u_{cc}(t)u_{ll}(t) - u_{cl}(t)u_{lc}(t)$.

The equations (12) and (14) are *compensated changes* in the individual demand functions. For example $\partial c(t)/\partial \omega(t)$ is the change in individual consumption when the after tax wage changes and the individual is compensated with initial capital so as to keep the marginal utility of capital at date t constant (i.e. keeping $q(t)$ constant). We see that compensated labour supply is increasing in the after-tax wage rate (equation (14)). The compensated cross-price effect depends on the sign of u_{lc} , i.e. whether marginal utility of consumption is increasing or decreasing with the amount of labour supplied. If utility is additively separable this term is zero (and the cross-price effect is zero). If consumption and leisure are complements, i.e. the marginal utility of consumption increases with leisure (and thereby decreases in labour), then $u_{lc} < 0$ and compensated consumption is decreasing in the after-tax wage rate. We see that (11) is negative if consumption is a normal good, and (13) is positive if leisure is a normal good. Typically, in dynamic economies sufficient for local stability of a steady state is that both consumption and leisure are normal goods.

4. Second-Best Optimal Taxation

In this section we shall solve for the second-best optimal tax programme, i.e. solving for the tax rates that give the highest utility to the representative individual subject to the (exogenous) public expenditure scheme under the assumption that the government can precommit to its optimal plan for the infinite future. Section 4.1 follows Chamley and derives the zero-capital income tax result and section 4.2 gives an interpretation. In section 4.3 we analyse the dynamic path of the capital-income tax for more general preferences than assumed in Chamley's original analysis. Next we prove in section 4.4 that the optimal labour income tax is positive at least from the date the non-confiscation constraint does not bind, implying that labour is taxed in steady state, and therefore under the second-best programme it is optimal to carry tax burden to the steady state. Section 4.4 analyses the dynamic paths of capital and labour taxes under different restrictions on preferences. Finally, section 4.5 employs a graphical analysis to study the effects on consumption, labour supply, and capital accumulation during the second-best tax programme.

4.1 Optimal Capital-Income Taxation: Chamley (1986)

The government chooses the time paths of $\tau^l(t)$ and $\tau^k(t)$ subject to the relevant constraints. The resource constraint could be obtained by subtracting the government's budget constraint from the individual's asset equation. Therefore one of these constraints is redundant in the optimisation (i.e. if two constraints are fulfilled, then the third is also fulfilled). We follow Chamley and take the resource constraint and the government's budget constraint as state equations. [It is convenient for the analysis since we may give an interpretation of the multipliers]. Next we have to take individual optimality into account, therefore we need to treat q as a state variable [see Kydland and Prescott (1980)].

We also know that it will be optimal to tax the factor which is inelastic as much as possible. Since individual's assets are inelastic at $t=0$, the government would want to tax away these assets. If we acknowledge private property rights we have to introduce a constraint on how much a government may tax assets. Following Chamley (1986) we assume that asset income may be taxed away, but not the assets themselves, and therefore we require the capital income tax to be less than or equal to 100%. This constraint is entirely arbitrary. If we instead had set the capital income tax equal to zero, we could have set the consumption tax "large" and combining with a "large" labour subsidy, which would have the effect of taxing the individual's initial assets (see Renström (1996a)). But here, again, how "large" the taxes are is also arbitrary. Only with an infinite consumption tax and an infinite labour subsidy we could "confiscate" the individual's initial assets.

Setting $\tau^c(t)=0$ and using CRS enable us to rewrite the government's budget constraint as

$$\begin{aligned}\dot{b}(t) &= \rho(t)b(t) + \rho(t)k(t) + \omega(t)l(t) - f(k(t), l(t)) + g(t) \\ b(0) &= b_0\end{aligned}\tag{15}$$

Dropping the time index and regarding c and l as functions of ω and q , we may write the current-value Hamiltonian as

$$H = u(c, l) + \mu \{ \rho \mathbf{b} + \rho k + \omega l - f(k, l) + \mathbf{g} \} + \lambda \{ f(k, l) - c - \mathbf{g} \} + \psi (\theta - \rho) q + v \rho \quad (16)$$

The necessary conditions for optimality are

$$H_\omega = 0 = u_c c_\omega + u_l l_\omega + \mu \{ l + \omega l_\omega - f_l(k, l) l_\omega \} + \lambda \{ f_l(k, l) l_\omega - c_\omega \} \quad (17)$$

$$H_\rho = 0 = \mu \{ \mathbf{b} + k \} - \psi q + v \quad (18)$$

$$H_b = \theta \mu - \dot{\mu} = \mu \rho \quad (19)$$

$$H_k = \theta \lambda - \dot{\lambda} = \mu \{ \rho - f_k(k, l) \} + \lambda f_k(k, l) \quad (20)$$

$$H_q = \theta \psi - \dot{\psi} = u_c c_q + u_l l_q + \mu \{ \omega l_q - f_l(k, l) l_q \} + \lambda \{ f_l(k, l) l_q - c_q \} + \psi (\theta - \rho) \quad (21)$$

These conditions are obtained by Chamley (1986). The zero capital-income tax result can be verified directly by inspection of equations (19)-(20). At steady state $\theta = \rho$ (by (19)) then (20) is $(\theta - f_k)(\lambda - \mu) = 0$, which can hold iff $f_k = \rho$.

4.2 Interpretation of the Zero Capital-Income Tax Result

Equation (19) tells us the following. μ is the marginal social value (marginal value to the social planner) of public debt, and thus is *negative* because an increase in public debt means that more revenue has to be raised by distortionary taxes (lump-sum taxes have been ruled out), and raising revenue through distortionary taxes is precisely the problem in this type of second best analysis. The marginal rate of substituting present tax burden for future tax burden, has to be contrasted with the marginal rate of transformation of present tax burden for future burden. That is, the social marginal rate of substitution of public debt between the present and the future has to equal the marginal rate of transformation between present and future debt. The $MRS(b(t+dt), b(t)) = \theta - \dot{\mu}(t)/\mu(t)$, and the rate of transformation is the *after-tax* interest rate ρ . Similarly λ is the social marginal value of capital. The $MRS(k(t+dt), k(t)) = \theta - \dot{\lambda}(t)/\lambda(t)$. The rate of transformation is capital's marginal product plus its marginal contribution to public funds, i.e. $(-\mu(t)/\lambda(t))(f_k - \rho)$. This is equation (20). Thus the marginal rate of substitution of tax burden can only equal the marginal rate of substitution of resources (capital) if capital is untaxed.

At a steady state the planner has to be indifferent transferring capital from today to the future (or vice versa), implying that the present marginal value of capital equals the future discounted marginal value of capital. At the steady state the planner is also indifferent of transferring tax burden from one date to the other, and therefore the present marginal value of public debt has to equal the future discounted marginal value of public debt. This means that their marginal rates of substitution have to be equal, and therefore capital has to be untaxed. If the government were to raise the capital tax at that moment,

it would find it more beneficial to transfer tax burden from today to the future (since $\rho < f_k$), by increasing public debt (or selling public assets). But raising the capital tax gives tax revenues, and since the labour tax is optimised, the capital tax has to adjust. Furthermore, if the capital tax is raised from zero, it is more beneficial transferring capital from today to the future, and the capital tax has to be lowered to accomplish that. So at the optimum those effects exactly cancel at the steady state.

We see that this result is implied by the definition of a steady state alone, and therefore the result applies for more general preferences than additively time separable. In fact, in the first part of Chamley (1986) the optimality of a zero-capital income tax in steady state was proven when the individual preferences are of the Koopmans (1960) form. However, the result generalises into an overlapping-generations economy when no such restriction is needed on individual utility functions, but where the social planner's social welfare function over generations is of the Koopmans form (see Renström (1998b)). This suggests that the zero capital-income tax result applies more generally than first was thought in the literature. Whenever an economy is characterised by preferences such that a steady state is possible, we should be able to verify the result (unless there are imperfections or more than one commodity is untaxable).

A natural question here is if the capital tax can reach zero out of the steady state, that is, if the marginal rate of substitution between present and future capital could equal the marginal rate of substitution between present and future public debt (or public assets). Since the marginal rates of substitutions derive from the underlying preferences, this would be a property of preferences alone. In the next section we shall see when the capital tax reaches zero in finite time.

4.3 The Path of the Optimal Capital-Income Tax

Chamley (1986) when analysing the dynamic path of the capital tax assumed a special class of preferences, additively separable in consumption and leisure and iso-elastic in consumption, i.e.

$$u(c, l) = [1 - \sigma]^{-1} c^{1-\sigma} + L(1-l) \quad (22)$$

However, we shall manipulate the necessary conditions slightly differently for general utility functions. Note that $u_c = q$ and $u_l = -q\omega$, and rewrite (17) and (21) to obtain

$$[q - \lambda]c_\omega + [-q\omega + \mu(\omega - f_l) + \lambda f_l]l_\omega + \mu l = 0 \quad (17')$$

$$\dot{\psi} = -[q - \lambda]c_q - [-q\omega + \mu(\omega - f_l) + \lambda f_l]l_q + \psi\rho \quad (21')$$

Substitute (17') into (21') to eliminate $[-q\omega + \mu(\omega - f_l) + \lambda f_l]$ and premultiply by q

$$\dot{\psi}q = [q - \lambda](c_\omega l_q q / l_\omega - c_q q) + \mu l l_q q / l_\omega + \psi\rho q \quad (23)$$

which by using (11)-(14) may be written as

$$\frac{d}{dt}(\Psi \mathbf{q}) = \Theta \Psi \mathbf{q} - [\mathbf{q} - \lambda] u_c / u_{cc} + \mu l \omega + \mu u_{lc} l / u_{cc} \quad (24)$$

Next we have

$$\frac{d}{dt}(\mu \mathbf{a}) = \dot{\mu} \mathbf{a} + \mu \dot{\mathbf{a}} = (\Theta - \rho) \mu \mathbf{a} + \mu(\rho \mathbf{a} + \omega l - c) = \Theta \mu \mathbf{a} + \mu \omega l - \mu c \quad (25)$$

Therefore, combining (24), (25), and (18) gives

$$\dot{v} = \Theta v - Z u_c / u_{cc} \quad (26)$$

where

$$Z \equiv \mathbf{q} - \lambda + \mu M \quad (27) \quad M \equiv -u_{cc} c / u_c - u_{lc} l / u_c > 0 \quad (28)$$

Taking the time derivative of Z gives

$$\dot{Z} = (\Theta - \rho) Z + (\lambda - \mu)(f_k - \rho) + \mu \dot{M} \quad (29)$$

When the utility function is of the form assumed by Chamley (i.e. equation (22)) then $M = \sigma$ and $\dot{M} = 0$. Chamley shows that the constraint $\rho \geq 0$ cannot be binding forever since the marginal utility would go to infinity if that was the case. Call the date at which $\rho \geq 0$ cease to bind t_1 . Then Chamley shows that $v(t_1) = \dot{v}(t_1) = Z(t_1) = 0$ and we have $\dot{Z}(t_1) = 0$. Finally (29) implies (by the choice of utility) that $[\lambda(t_1) - \mu(t_1)][f_k(t_1) - \rho(t_1)] = 0$, which can hold iff $f_k(t_1) = \rho(t_1)$, i.e. $\tau^k(t_1) = 0$. Thus there is a regime switch at t_1 . Before that date the capital income tax is 100%, and thereafter the tax is zero.

This simple dynamics for the capital tax can be obtained for a class of utility functions more general than the one employed by Chamley. To see this we may proceed as follows. When the constraint on the capital income tax does not bind, $v = \dot{v} = 0$, we have $Z = \dot{Z} = 0$, so (29) gives

$$-\mu^{-1}(\lambda - \mu)(f_k - \rho) = \dot{M} \quad (30)$$

So the capital tax is zero if and only if M is constant. For this to happen out of steady state we need equation (28) to be constant $\forall t$. Integrating (28) we obtain the class of utility functions that can be expressed as

$$u_c(c, l) = c^\gamma l^\gamma \Phi(c/l) \quad (31)$$

where γ is a positive constant and $\Phi(\cdot) > 0$ is an arbitrary function. Thus we see that for all other utility functions when the confiscation constraint ceases to bind, the capital-income does not jump to zero. We shall now turn to the optimal labour income tax.

4.4 The Optimal Labour-Income Tax

The labour income tax has previously been ignored, leaving the possibility open that all taxes could be zero at the steady state, meaning that the government accumulates assets during the initial period of the tax programme, and uses these assets to fund the its

expenditure in the future. However, in the second best, for the economy we have described, this is not the case.

Theorem 1 *Under the second-best tax programme, sufficient for the labour income tax being positive at least from the date when capital income is not confiscated, and onwards, is that leisure is a normal good.*

Proof: Rewrite (17') as

$$-\mu^{-1}(\lambda - \mu)(f_l - \omega) = l/l_\omega - \mu^{-1}(q - \lambda)(\omega - c_\omega/l_\omega) \quad (32)$$

When $Z=0$ we have $-(q-\lambda)/\mu = M$ which implies that $q > \lambda$ when $\rho \geq 0$ does not bind. Since $\omega - c_\omega/l_\omega = l_q q/l_\omega > 0$ (by normality of leisure) and $q > \lambda$, the right-hand side of (32) is positive. Since $\lambda > 0$ and $\mu < 0$ the left-hand side of (32) is positive iff $f_l > \omega$. QED

This means that it is optimal to carry tax burden at all dates, and intuitively is related to the normality in consumption goods at different dates (implied by the additively time-separable utility). The nature of the second best implies that the consumption possibilities are smaller than in the first-best, and with normality in consumption goods it is optimal to reduce consumption at all dates, i.e. distort consumption at all dates. We should notice that we refer to normality of consumption across dates.

4.5 Transition Paths of Optimal Taxes

We shall analyse the transition paths of the labour tax as well as of the capital tax, under different restrictions on the instantaneous utility function. First we shall assume additive separability in consumption and leisure, then denoting $\eta(t) \equiv \omega l_\omega/l$ equation (32) becomes

$$\tau^l/(1 - \tau^l) = -\eta^{-1}\mu/(\lambda - \mu) + (q - \lambda)/(\lambda - \mu) \quad (33)$$

The multipliers are functions of time. Combining (19) and (20) and integrating between t_0 and t we obtain

$$\lambda(t) - \mu(t) = [\lambda(t_0) - \mu(t_0)] e^{\int_{t_0}^t (\theta - f_k(s)) ds} \quad (34)$$

Similarly integrating (19) and (9) and substituting for $q(t)$, $\lambda(t)$ and $\mu(t)$ in (33) gives us

$$\frac{\tau^l}{1 - \tau^l} = \left[\frac{-\mu(t_0)}{\lambda(t_0) - \mu(t_0)} \frac{1}{\eta} + q(t_0) - \mu(t_0) \right] e^{\int_{t_0}^t (f_k(s) - \rho(s)) ds} - 1 \quad (35)$$

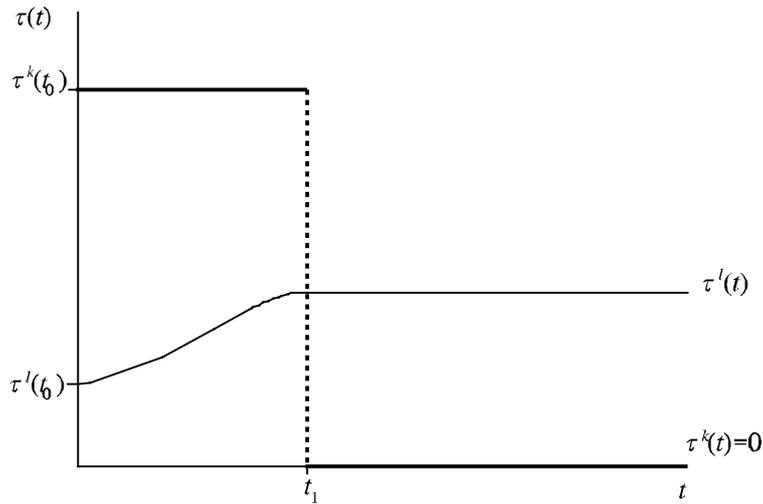
We begin by analysing the policy when utility is iso-elastic both in consumption and in labour supply

$$u = [1 - \sigma]^{-1} c^{1-\sigma} - [1 + 1/\eta]^{-1} l^{1+1/\eta} \quad (36)$$

We know from section 4.3 (and from Chamley's analysis) that when the utility function is additively separable in consumption and leisure and iso-elastic in consumption, in the beginning when the optimal programme is implemented (at t_0 say), the capital income tax is 100% (the assumed upper limit). After finite time (at t_1 say) there is a regime switch, and the capital income tax becomes zero.

The iso-elasticity in labour supply gives $\omega l_\omega/l = \eta = \text{constant}$, and the dynamics of $\tau^l(t)$ depends *only* on the integral in (35). Then, since the bracketed term in (35) is positive we see that $\tau^l(t)$ is an increasing function of time if $f_k > \rho$, i.e. if the capital income tax is positive. So, between t_0 and t_1 the labour income tax is increasing over time. After t_1 the labour income tax is constant (since the capital income tax is zero, due to the iso-elastic function of consumption). The dynamic paths for the optimal labour- and capital-income taxes, for utility of the form (36), are depicted below.

Figure 1

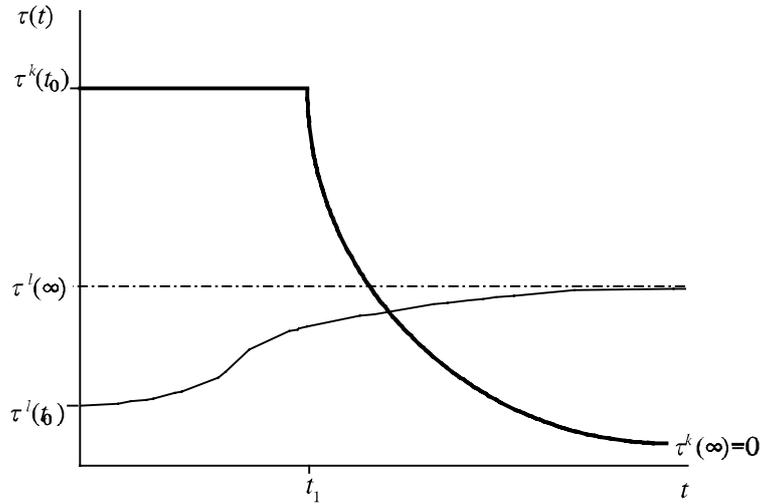


The dynamic path for the capital income tax in Chamley's economy is very similar to the dynamics of controls in *most-rapid-approach* solutions (see Kamien and Schwartz (1991) pp. 97-101). They should not be mixed up, however. A most-rapid-approach path is usually found in economies with linear objectives (such as linear instantaneous utility of consumption). Such an optimisation problem would have a corner solution (such as consume as much as possible, or as little as possible). This control would keep its extreme value until the state reaches some particular value (e.g. in a consumption-savings economy, when the capital stock reaches its steady state value). The concept of most-rapid-approach paths implies that a particular value of the state should be attained as "quickly as possible." This interpretation is not valid in the Chamley economy. In the latter it is not the case that we wish to "tax as much as possible" until, say, we have no further revenue requirement,

since labour is taxed at steady state.

As we noted in section 4.3 it is only when utility is separable in consumption and labour supply and iso-elastic in consumption that the capital income tax is either as large as possible (say 100%) or zero. For all other utility specifications it is optimal to adjust the capital income tax gradually towards zero after the constraint cease to bind (provided that the economy actually converges to a steady state). Consequently, even if utility is iso-elastic in *labour*, the labour income tax does not become constant at t_1 , but continues to increase towards its steady-state value. This case is depicted below.

Figure 2



We shall now turn to an analysis of adjustment paths of capital, consumption and labour supply under the second-best tax programme.

4.6 Economic Adjustment under the Optimal-Tax Programme

We shall graphically analyze the dynamic paths of capital, consumption and labour supply under the optimal policy. The evolution of the capital stock and the private co-state may be written in terms of the private co-state and the labour income tax

$$\dot{k}(t) = f(k(t), l(q(t), \omega(t))) - c(q(t), \omega(t)) - g(t) \quad (37)$$

$$\dot{q}(t) = [\theta - f_k(k(t), l(q(t), \omega(t)))]q(t) \quad (38)$$

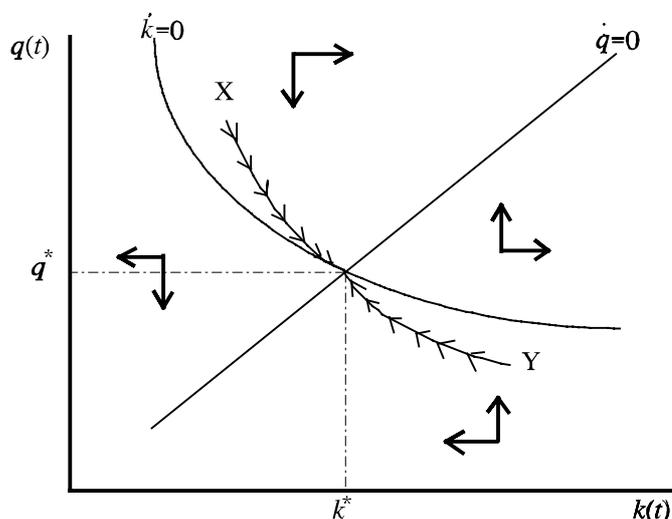
where $\omega(t) = [1 - \tau^l(t)]f_l(k(t), l(q(t), \omega(t)))$.

Setting $\dot{k} = 0$ in (37) gives us all combinations of q and k consistent with a constant capital stock. We view these combinations as a functional relationship between q and k . It is plausible that q is a decreasing convex function of k .⁷ This function is depicted

graphically in Figure 3 below. For a capital stock to the left of this curve, capital is decreasing because, for a given q , consumption is too high relative to production to maintain the current capital stock, and therefore q would need to be higher (= lower consumption) to compensate for this. The converse is true to the right of the curve.

Similarly, setting $\dot{q} = 0$ in (38) gives us a functional relationship between q and k , consistent with a constant q . It is plausible that q is an increasing function of k .⁸ This function is depicted as a line in Figure 3 below. For capital to the left of this line, the capital stock is smaller than the quantity consistent with $\rho = \theta$. For a given q this implies that $(1-\tau^k)f_k$ is greater than θ , (f is concave in k), and in turn that q is decreasing. The converse is true to the right of the line. Together the line and the curve from the usual saddle-path diagram.⁹

Figure 3



If the functions are "well-behaved" we have a unique steady state which is at least locally stable. Global stability may be obtained by imposing restrictions on utility and production functions so as to rule out the unstable paths as sub-optimal. The above diagram is helpful in analysing the out-of-steady-state dynamics under the optimal tax policy. We will take (k^*, q^*) to be the steady state under the optimal tax policy.

It is instructive to concentrate on iso-elastic utility, since then we have two regimes for the after-tax interest rate: $\rho = 0$ and $\rho = f_k$. We will see later on that the dynamic behaviour of the economy does not fundamentally change for the more general case. Also, for expositional simplicity we keep τ^l constant in the graphs. We will see later on that when τ^l is changing over time (according to the optimal policy), the fundamental dynamic patterns do not change. We shall consider the following timing: at t_0 the optimal policy is implemented, and at t_1 the capital-income tax is set to zero.

We may think of three initial values of the capital stock when this policy is implemented: lower, equal to, or greater than its steady-state value. That is, either $k(t_0) < k^*$, or $k(t_0) = k^*$, or $k(t_0) > k^*$. Assume first that the initial value of the capital stock is equal

to its long-run value. When $\rho = 0$ both the slope and the level of the line $\dot{q} = 0$ goes to plus infinity, (we may think of this as when the line disappears to the left), and we are left with only the curve, as in Figure 4 below.

We then have three possibilities for the individual's behaviour. The individual may choose initial consumption and initial labour supply so that initial marginal utility is

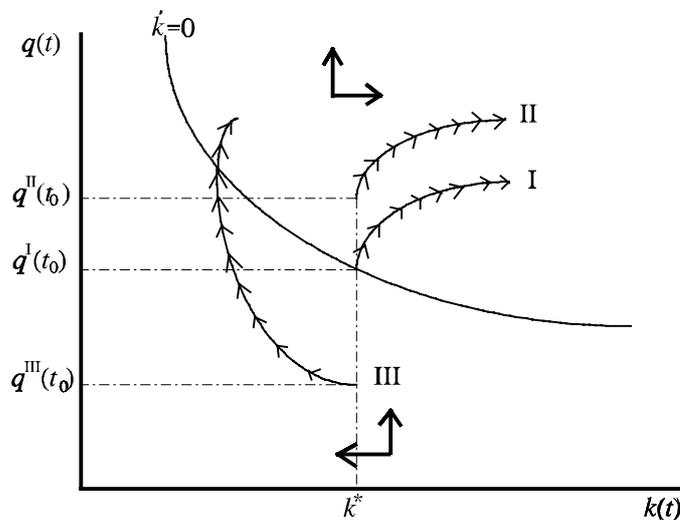
(I) equal to steady-state marginal utility, i.e. $q(t_0) = q^*$,

(II) greater than the steady-state level, i.e. $q(t_0) > q^*$,

(III) smaller than the steady-state level, i.e. $q(t_0) < q^*$,

and then adjusting according to the first-order differential equation $\dot{q} = \theta q$. Graphically the three possibilities are depicted in Figure 4, below.

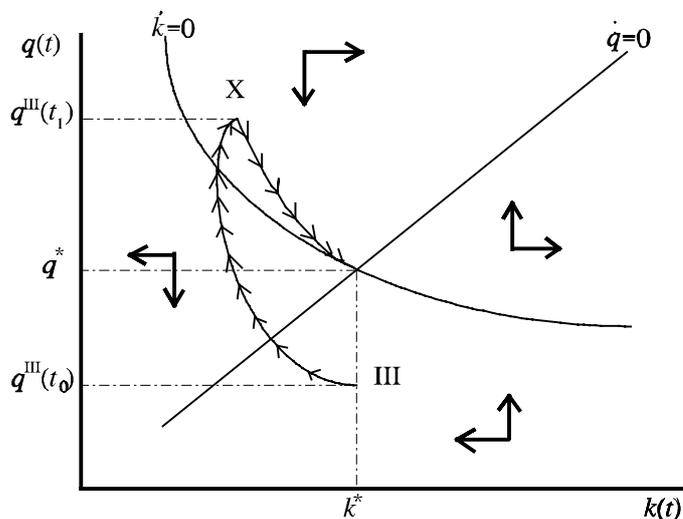
Figure 4



It is clear that (I) cannot be optimal, when the regime switches to a zero capital income tax the economy cannot turn back to its steady state. The economy would accumulate capital forever and consumption would go to zero. On the same grounds we can rule out (II). We are left with (III) as the only possibility. In fact the initial value of q is such that the economy is guaranteed to join the unique converging trajectory X, exactly at the date of the regime switch t_1 . This path requires capital decumulation.¹⁰ See Figure 5, below.

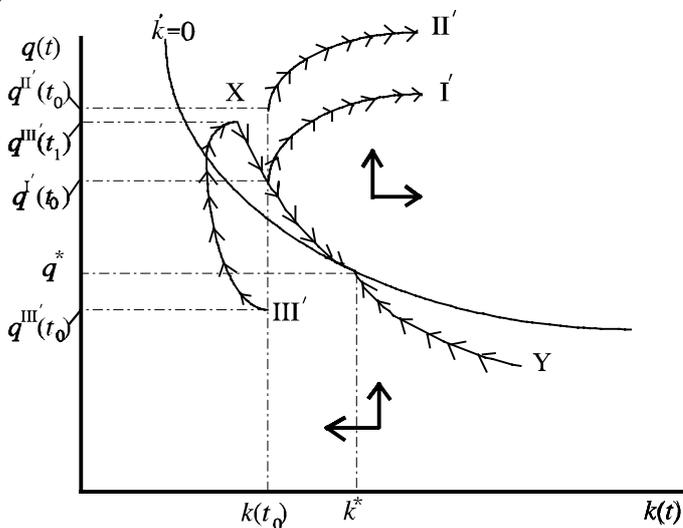
It is clear that the analysis above does not fundamentally change when there is no regime switch but a gradual adjustment of ρ towards f_k , (i.e. when form of the utility function is more general than the one in equation (31)). Still the economy would decumulate capital for a period and join the new converging trajectory associated with the continuous adjustment of ρ towards f_k . The same analysis also applies when the economy's initial capital stock is different from its steady-state value. To see this suppose the initial capital stock is smaller than its steady state value.

Figure 5



We then have three possibilities analogously to the case when $k_0 = k^*$: The individual may choose initial marginal utility (I') on trajectory X, (II') above trajectory X, or (III') below trajectory X. Case (I') and (II') may be ruled out on the same grounds as above. We are left with (III'), which at t_1 joins X. See figure 6.

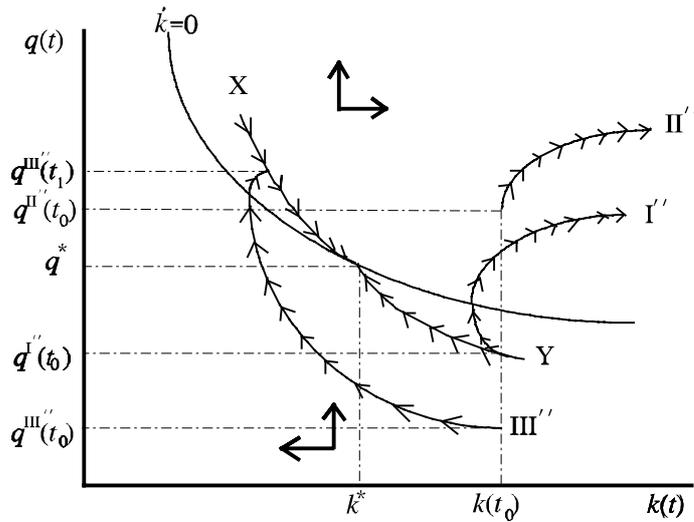
Figure 6



Finally suppose that the economy's capital stock is greater than its steady state value. Again we have three possibilities. The individual may choose a q on trajectory Y, and then follow the differential equation $\dot{q} = \theta q$. But then the economy cannot follow trajectory Y,

since q grows faster when $\rho = 0$ than when $\rho = f_k$. Since q grows faster, consumption will decrease faster and k will not decumulate as quickly as intended on trajectory Y. Clearly this implies that the economy will go on a trajectory above, and after some time cross the $\dot{k} = 0$ line and behave as in the case (I) above. By the same reason we may rule out any $q(t_0)$ above the trajectory and we are left with case (III''), depicted in Figure 7. As above the economy will decumulate capital.

Figure 7



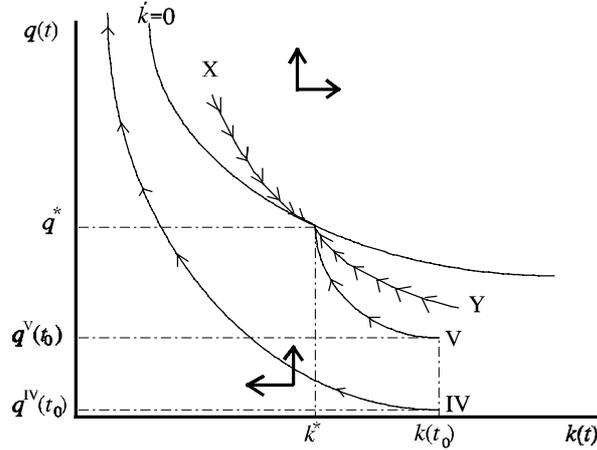
The length of the confiscatory regime ($t_1 - t_0$) depends on the economy's funding requirement. If the funding requirement is large relative to the capital stock $k(t_0)$ then the co-state $q(t_0)$ would take on a *lower* value and join X at a later date. So, the greater the distance $|t_1 - t_0|$ the smaller the $q(t_0)$. Smaller $q(t_0)$ is associated with faster capital decumulation.

We may think of two paths not covered by the Figures 3-5. First, if the confiscatory regime is permanent, i.e. $|t_1 - t_0| \rightarrow \infty$, then $q(t_0)$ takes on a value so small that the trajectory never crosses the $\dot{k} = 0$ line and therefore continues to infinity and the capital stock goes to zero. Second, if the funding requirement is small *and* $k(t_0) > k^*$, we have the possibility that $q(t_0)$ is large enough for the economy not to decumulate capital below k^* . The two cases are depicted below in Figure 8, as trajectories IV and V respectively.

We have drawn the graphs for a *constant* labour income tax τ^l . When τ^l changes over time we have to imagine the $\dot{k} = 0$ line "moving" over time (at least between t_0 and t_1) in Figures 4, 6-8. For example, when τ^l is increasing over time [as for the utility function in (36)] the $\dot{k} = 0$ line will move outwards as time goes on. This is so since an (uncompensated) increase in the labour income tax decreases labour supply and at least not decreases consumption, if consumption and leisure are complements. This would make $\dot{k} < 0$. To "restore" $\dot{k} = 0$ the level of k has to be higher (or alternatively the level of q has to be greater). If this is the case, trajectory III will as before be below the $\dot{k} = 0$ line, and towards t_1 "chase" the line and cross it. Clearly, the fundamental pattern of trajectory III

does not change, and on the same basis as before we can always rule out the trajectories I and II.

Figure 8



From the graphical analysis we have found a common characteristic of the economy under the confiscatory regime:

Remark *Regardless the initial value of the capital stock when the second-best policy is implemented, the economy always decumulates capital in the beginning of the regime, and at least after some time (if the confiscatory regime is long enough) the pre-tax interest rate becomes greater than the rate of time preference, and the economy grows toward its new steady state (accumulating capital).*

We shall proceed with the analysis by analysing the case with heterogeneous individuals in the next section.

5. Heterogeneous Individuals

We shall see that the zero-capital income tax carries over to an economy with heterogeneous individuals. Judd (1985) has proved this for a two-individual case with a Paretian welfare function (a weighted average of the two individuals), for preferences that are not necessarily additively time separable, but separable of the Usawa (1968) form.

We shall proceed in a different way. We shall ask an arbitrary individual to state her most preferred tax sequences. This may be interpreted, at this stage, as either a Rawlsian welfare function (where the worst off individual chooses taxes) or a representative democracy where the median individual chooses tax policy (but under full precommitment). We will assume that individuals differ linearly in wages, so the pre-tax wage of individual i is $w^i(t) = \gamma^i w(t)$, where γ^i is the productivity parameter, which is normalised such that the population average equals unity: $\int \gamma^i dF(i) = 1$. Each individual's budget constraint is

$$\dot{k}^i(t) = \rho(t)k^i(t) + \omega^i(t)l^i(t) - c^i(t) + g(t), \quad k^i(0) = k_0^i \quad (39)$$

where $g(t)$ is a lump-sum transfer at date t , equal for all individuals. The average capital stock and the average labour supply (in efficiency units) are $\bar{k}(t) \equiv \int k^i(t) dF(i)$ and $\bar{l}(t) \equiv \int l^i(t) dF(i)$ respectively. Because of the lump sum transfer and the infinitely lived individuals this economy is Public Debt Neutral. This means that the behaviour of the economy, and the utilities of the individuals, are the same even if public debt was allowed. Without loss of generality we therefore impose period-by-period government budget balance by setting $b(t)$ and its time derivative to zero in equation (4).

Next, using (4) in individual i 's budget constraint (39) to eliminate $g(t)$ gives

$$\dot{\Delta}^i(t) = \rho(t) \Delta^i(t) + \omega^i(t) (\gamma^i l^i(t) - \bar{l}(t)) \quad (40)$$

where $\Delta^i(t) \equiv k^i(t) - \bar{k}(t)$ is the difference between the capital holding of individual i and the average (aggregate) capital stock. An individual solving (1) subject to (40) (and taking the aggregate quantities as given) gives the optimality conditions (7)-(9), evaluated at individual quantities and individual wage. The Hamiltonian for the decisive individual's problem is

$$H^i = u(c^i, l^i) + q^i \{ \rho [k^i - \bar{k}] + \omega [\gamma^i l^i - \bar{l}] + f(\bar{k}, \bar{l}) - c^i \} + \lambda \{ f(\bar{k}, \bar{l}) - \bar{c} \} + \int \psi^i [\theta - \rho] q^i dF(i) + v \rho \quad (41)$$

The optimality conditions with respect to $k^i(t)$ and $\bar{k}(t)$ are respectively

$$H_{k^i(t)} = \theta q^i(t) - \dot{q}^i(t) = q^i(t) \rho(t) \quad (42)$$

$$H_{\bar{k}(t)} = \theta \lambda - \dot{\lambda}(t) = q^i(t) [r(t) - \rho(t)] + \lambda(t) r(t) \quad (43)$$

Equation (43) is the same as equation (20), and (42) has the same structure as (19), thus the arbitrary individual prefers zero capital tax at the steady state. Thus there is unanimity on zero capital-income taxation. Not the most extreme welfare function (such as Rawls) can change this result!

Theorem 2 *Assume purely redistributive taxation and heterogeneous individuals, then if an individual can choose capital-income and labour-income taxes for the entire future and if the economy under these tax paths goes to a steady state the preferred capital-income tax is zero in this steady state, regardless who the decisive individual is.*

The labour tax takes a very similar form as in the single-individual economy, if we make an additional assumption regarding individuals' preferences, such that aggregation occurs (i.e. the aggregate economic equilibrium being independent of distributional characteristics). If we assume that individuals preferences are either additively separable in consumption and leisure, or multiplicatively separable. Then, necessary and sufficient for aggregation is that the sub-utility functions belong to the HARA-family (Hyperbolic Absolute Risk Aversion). Then the labour tax preferred by individual i is (see Renström (1997) for details)

$$\frac{\tau^l(t)}{1-\tau^l(t)} = \left[\bar{l}(t) - \gamma^l l^i(t) \right] \left(\frac{\partial \bar{l}(t)}{\partial \omega(t)} \omega(t) \right)^{-1} + \tilde{\delta} \left[\frac{q^i(0)}{q^i(0)+\lambda(0)} e^{\int_0^t [r(s) - \rho(s)] ds} - \tilde{\delta} \right] \quad (44)$$

where $\tilde{\delta} = \{1 \text{ if } u_{c^l}=0, \delta \text{ if } u_{c^l} \neq 0\}$. δ is the parameter in the (HARA) sub-utility function, where consumption enters as an argument, that takes on positive value for Decreasing ARA, zero for Constant ARA, and negative for Increasing ARA. Comparing the labour tax in (44) with the one in the single-individual framework (35), we see that they take the same form. However, here labour is taxed only if the labour supply in efficiency units is smaller than the average labour supply, for then the decisive individual gains from the redistributive labour tax. For iso-elastic disutility of labour, $\omega_{\omega}^l/l = \eta = \text{constant}$, the dynamic path of the redistributive labour tax is analogous to Figures 1-2.

To conclude, the second best tax structure in single individual economies coincides quite well with the second best tax structure in heterogeneous individual economies.

6. Third-Best Optimal Taxation

6.1 The Optimal Tax Programme without Precommitment

It is quite intuitive what would happen in the economy described in section 2, if the government could not precommit to future policy. At each date the non-confiscation constraint would be binding and the capital-income tax would be 100% as long as the government has a funding requirement. The government would then accumulate assets large enough to cover all future expenditures. The steady state is characterised by all taxes being zero. We have a *zero-tax equilibrium*, no tax burden is carried in steady state. This is formally proven in Renström (1997) by modelling the time-consistent tax problem as a differential game. There is another potential steady state, where the individual owns zero assets, however it is shown that this steady state cannot be attained.

We have a rather paradoxical situation, which gives us a word of *warning* when comparing second-best and third-best taxation. In the third best the economy attains a steady state with higher welfare than in the second best! Therefore, evaluation of any partial precommitment mechanism by comparing the steady states only would give misleading results. Any partial precommitment mechanism which makes the economy closer to the second best, will result in a steady state with tax burden, and therefore the steady state welfare will be lower than in the third best.

6.2 Partial Precommitment Solutions

Ruling out instruments

The problem in the third-best is the overtaxation of capital in the beginning of the optimisation period. One could then think that ruling out the capital tax in the constitution could partly solve the problem. However, taxing of capital is possible with a consumption tax and a labour subsidy, because of tax equivalence (see Renström (1996a)). So even ruling out an instrument does not solve the problem.

Timing

A possibility is a constitution under which the taxes have to be decided upon well in advance of their implementation (i.e. constitutional delays). This has been explored in Krusell and Ríos-Rull (1994), Krusell, Quadrini and Ríos-Rull (1996), and in Renström (1998*b*). Another possibility is to reduce the frequency of policy revisions. So if the government, when it takes a fiscal decision, it has to live with it longer. During the time interval when policy cannot be changed individuals would decumulate capital. This is very sensitive to the fiscal structure. In a representative individual economy with access to consumption taxes and labour subsidies, the problem of no distortionary taxation at the steady state comes back (Renström (1997)). However, it does seem to work with a partial set of redistributive taxes (Krusell and Ríos-Rull (1994), Krusell et.al. (1996)).

Evasion

If the government cannot costlessly observe income, individuals have a possibility of evading taxes. However, the government could reduce tax evasion considerably by letting the punishment for evasion getting sufficiently large. But the government may find it optimal to strategically choose low punishment for tax evasion and low auditing rates, so that tax evasion occurs in equilibrium. Tax evasion makes the tax base more elastic and the tax program is then closer to the second best. This has been independently analysed in Boadway and Keen (1998) and in Renström (1998*a*).

Tax Earmarking

Earmarking of tax revenue can also improve upon the third best, if the earmarking rules are chosen before the fiscal decision is taken (or alternatively if earmarking rules are chosen less often than the taxes are decided upon). Such constitutional earmarking is explored in Renström (1998*c*) and in Marsiliani and Renström (1998). There it is shown that the rules can get the third-best equilibrium closer to the second-best, but not completely coincide with the second-best equilibrium. The reason is that tax earmarking simply is a partial precommitment device.

Interestingly, these constraints would be sub-optimal in the static optimal tax literature (since the third best never arises there), but may give a welfare improvement in a dynamic economy, if only the third best is possible.

7. Conclusions

In this paper we have reviewed the recent literature on dynamic taxation, and carefully explored the most important findings regarding capital and labour taxes. We have also showed when the results from single individual analysis generalise to an economy with heterogeneous individuals. We have pointed out that, since the second best tax programme is time inconsistent, one of the most challenging task for the researcher is to propose constraints on governments in such a way that the time-consistent policy comes closer to the second best. This is a rather unexplored area, and it is probably here that most attention will be directed in the future.

References

- Atkinson, Anthony B., and Agnar Sandmo (1980), "Welfare Implications of the Taxation of Savings," *Economic Journal*, 90: 529-549.
- Atkinson, Anthony B., and Joseph E. Stiglitz (1972), "The Structure of Indirect Taxation and Economic Efficiency," *Journal of Public Economics*, 1: 97-119.
- Barro, Robert J. (1974), "Are Government Bonds Net Wealth?," *Journal of Political Economy*, 82(6): 1095-1117.
- Başar, Tamer, and Geert J. Olsder (1982), *Dynamic Noncooperative Game Theory*, Academic Press, London/New York.
- Blanchard, Olivier J., and Stanley Fischer (1989), *Lectures on Macroeconomics*, MIT Press, Cambridge, Massachusetts, US.
- Boadway, Robin, and Michael Keen (1998), "Evasion and Time Consistency in the Taxation of Capital Income," *International Economic Review*, 39(2): 461-476.
- Chamley, Christophe (1985), "Efficient Taxation in a Stylized Model of Intertemporal General Equilibrium," *International Economic Review*, 26(2): 451-468.
- Chamley, Christophe (1986), "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, 54(3): 607-622.
- Cohen, Daniel, and Philippe Michel (1988), "How Should Control-Theory be Used to Calculate a Time Consistent Government Policy," *Review of Economic Studies*, 55(2): 263-274.
- Correia, Isabel H. (1996), "Should Capital Income be in the Taxed Steady State?," *Journal of Public Economics*, 60(1): 147-151.
- Diamond, Peter A. (1965), "National Debt in a Neo-Classical Growth Model," *Journal of Political Economy*, 55: 1126-1150.
- Diamond, Peter A. (1973), "Taxation and Public Production in a Growth Setting," in James A. Mirrlees and Nicolas H. Stern eds., *Models of Economic Growth*, MacMillan, London.
- Diamond, Peter A. (1975), "A Many-Person Ramsey Tax Rule," *Journal of Public Economics*, 4: 227-244.
- Diamond, Peter A., and James A. Mirrlees (1971a), "Optimal Taxation and Public Production I: Production Efficiency," *American Economic Review*, 61: 8-27.
- Diamond, Peter A., and James A. Mirrlees (1971b), "Optimal Taxation and Public Production II: Tax Rules," *American Economic Review*, 61: 261-278.
- Jones, Larry E., Rodolfo E. Manuelli, and Peter E. Rossi (1997), "On the Optimal Taxation of Capital Income," *Journal of Economic Theory*, 73: 93-117.
- Judd, Kenneth L. (1985), "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, 28: 59-83.
- Kamien, Morton I., and Nancy L. Schwartz (1991), *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*, North-Holland, Amsterdam, London.
- Kemp, Murray C., Ngo Van Long, and Koji Shimomura (1993), "Cyclical and Noncyclical Redistributive Taxation," *International Economic Review*, 34(2): 415-429.
- Koopmans, Tjalling C. (1960), "Stationary Ordinal Utility and Impatience," *Econometrica*, 28(2): 287-309.

- Krusell, Per, and José-Víctor Ríos-Rull (1994), "What Constitutions Promote Capital Accumulation? A Political Economy Approach," Working Paper No. 1, W. Allen Wallis Institute of Political Economy, University of Rochester.
- Krusell, Per, Vincenzo Quadrini, and José-Víctor Ríos-Rull (1996), "Are Consumption Taxes Really Better than Income Taxes?," *Journal of Monetary Economics*, 37: 475-503.
- Kydland, Finn E., and Edward C. Prescott (1977), "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, 85(3): 473-491.
- Kydland, Finn E., and Edward C. Prescott (1980), "Dynamic Optimal Taxation, Rational Expectations and Optimal Control," *Journal of Economic Dynamics and Control*, 2: 79-91.
- Lucas, Robert E., Jr., and Nancy L. Stokey (1983), "Optimal Fiscal and Monetary Policy in an Economy without Capital," *Journal of Monetary Economics*, 12: 55-93.
- Lucas, Robert E., Jr. (1990), "Supply-Side Economics: An Analytical Review," *Oxford Economic Papers*, 42: 293-316.
- Marsiliani, Laura, and Thomas I. Renström (1998), "Time-Inconsistency in Environmental Policy: Tax Earmarking as a Precommitment Solution," mimeo, Tilburg University.
- Mirrlees, James A. (1971), "An Exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies*, 38: 175-208.
- Myles, Gareth D. (1995), *Public Economics*, Cambridge University Press.
- Ordober, Janus A., and Edmund S. Phelps (1979), "The Concept of Optimal Taxation in the Overlapping-Generations Model of Capital and Wealth," *Journal of Public Economics*, 12: 1-26.
- Park, No-Ho (1991), "Steady-State Solutions of Optimal Tax Mixes in an Overlapping-Generations Model," *Journal of Public Economics*, 46: 227-246.
- Ramsey, Frank P. (1927), "A Contribution to the Theory of Taxation," *Economic Journal*, 37: 47-61.
- Ramsey, Frank P. (1928), "A Mathematical Theory of Saving," *Economic Journal*, 38: 359-559.
- Renström, Thomas I. (1996a), "Dynamic Taxation and Equivalent Tax Systems," Department of Economics Discussion Paper No. 96-06, The University of Birmingham.
- Renström, Thomas I. (1996b), "Endogenous Taxation: An Overlapping Generations Approach," *Economic Journal*, 106: 471-482.
- Renström, Thomas I. (1997), "Endogenous Taxation in a Dynamic Economy," D.Phil Thesis, University of Oxford.
- Renström, Thomas I. (1998a), "Tax Evasion as a Disciplinary Mechanism for Fiscal Policy," in Shibata, H., and T. Ichori (Eds.), *The Welfare State, Public Investment, and Growth*, Springer-Verlag, Tokyo.
- Renström, Thomas I. (1998b), "On Optimal Taxation in Overlapping Generations Economies," mimeo, Tilburg University.
- Renström, Thomas I. (1998c), "Dynamic Tax Earmarking and Endogenous Fiscal Structure," mimeo, University of Birmingham.
- Turnovsky, Stephen J. (1995), *Methods of Macroeconomic Dynamics*, MIT Press, Cambridge, Massachusetts, US.

Usawa, Hirofumi (1968), "Time Preference, the Consumption Function and Optimum Asset Holdings," in J.N. Wolfe (ed.), *Value, Capital and Growth: Papers in Honour of Sir John Hicks*, University of Edinburgh Press.

Endnotes

¹ The Ramsey rule would only tell us that we should induce the same proportionate changes in compensated demand on all commodities, the tax may then be larger on commodities which are less price elastic.

² See Myles (1995) for a review.

³ By the representative-consumer assumption we ignore the equity aspect, and also the motivation why the tax system is distortionary in the first place!

⁴ One exception is Lucas and Stokey (1983) who show how time consistent policy may be sustained in an economy without capital, if the government can choose the maturity structure of public debt.

⁵ The methodology is set out in Başar and Olsder (1982). Another relevant paper is Cohen and Michel (1988) which explores the different solution concepts in an economy with hypothetical quadratic loss functions and linear constraints.

⁶ Despite this it is not automatic that the First and Second Welfare Theorems apply in *absence* of distortionary taxes. It is well known that the OLG economies can give rise to dynamic inefficiency (if population growth is "large").

⁷ Decreasing is quite obvious. l is positively related to q and c is negatively related to q , then a high value of q in (37) requires k to be lower (to maintain $\dot{k} = 0$). Convexity can be verified.

⁸ Increasing because, a greater value of k implies a lower f_k (since $f_{kk} < 0$). Then, to keep $f_k = \theta$, we require l to be greater (since $f_{kl} > 0$) and therefore q to be greater.

⁹ Saddle path diagrams in the literature are usually written in the consumption - capital space (c and k), [e.g. Blanchard and Fischer (1989)]. However, since we have an economy with endogenous labour supply, it is more convenient to use the costate q instead of c and l , [e.g. Turnovsky (1995)]. We can always translate from q to consumption and labour supply respectively.

¹⁰ We could think of a fourth possibility as well: $q(t_0)$ is very low. Then the trajectory would not cross the $\dot{k} = 0$ line, but continue left upwards toward the q -axis. In this case marginal utility goes to infinity and the economy's capital stock goes to zero. This would happen if the $\rho = 0$ regime lasts forever.