ON OPTIMAL TAXATION IN
OVERLAPPING GENERATIONS ECONOMIES*

Thomas I. Renström

This version, August 2000

The paper studies optimal taxation in an overlapping generations economy, assuming different degrees of commitment: a) full commitment (second best), b) no commitment, c) partial commitment (next periods’ taxes are chosen today). It is shown that the second best optimal capital income tax is zero in steady state if the social welfare function is weakly separable across generations. That steady state also obeys the Modified Golden Rule. Third best optimal policy is also solved for (case b) and c)). In case b) the optimal capital income tax is generally not zero, and the steady state is generally not the Modified Golden Rule. For special case of preferences (where individuals’ utility functions are logarithmic), the steady state is the Modified Golden Rule, and furthermore the no commitment tax policy coincides with the tax policy in the full commitment case. In case c) the capital income tax is generally not zero, but the steady state is the Modified Golden Rule. We also derive the class of individual utility functions that are necessary and sufficient for the capital tax being zero at all dates in cases a) and c).

Keywords: Optimal taxation, fiscal policy, overlapping generations.

JEL classification: H21, C73, D91, E62.

*I am grateful to James Mirrlees and Leslie Reinhorn for helpful comments.
1. INTRODUCTION

In this paper we solve for the optimal taxes in an overlapping-generations (OLG) economy, in the second best (full commitment) and in the third best (partial or no commitment). We find that the second best optimal capital income tax is zero in steady state if the social welfare function is recursively separable across generations (Koopmans (1960) form), even though individuals are heterogeneous, and even though individuals’ utility functions need not be separable. This result is in contrast with previous studies of OLG economies who have found the optimal capital income tax to be zero only in a special case: when individuals’ preferences are logarithmic in consumption and leisure.¹ The reason for the difference between previous studies and this paper is that the previous studies have not found the second best optimal policy, but the third best, by letting the government only choosing prices that affect one generation at a time.

The intuition for the zero-capital income tax result has little to do with uniform commodity taxation. Rather the result comes from the fact that at a steady state the government has to be indifferent along two margins of variation. First, the government has to be indifferent in transferring physical resources from generation date to the other, second the government has to be indifferent in transferring tax burden from one generation to the other. The rate of transformation of capital from one period to the other is the pre-tax interest rate, while the rate of transformation of tax burden (public debt) from one period to the other is the after-tax interest rate. The government can only be indifferent in both dimensions if capital is untaxed, for then the two stock’s rate of transformation are the same. This is an intergenerational optimality condition. It has nothing to do with the individuals within a generation being at a steady state, since a steady state in an OLG economy very well can be consistent with the individuals themselves having rising or declining consumption profiles over their life times.

There are also situations in which one wish not to tax capital at all, not even out of steady state. This happens when uniform commodity taxation is optimal. Taxing dated goods at the same rate translates into a zero capital-income tax. In this paper we will derive the

¹ One has therefore been tempted to believe that there has been a fundamental difference between the optimal tax programme in an OLG economy and in a dynastic economy, since it is a general result that the optimal capital income tax is zero in a dynastic economy with individual’s preferences being of the Koopmans form [Chamley (1986)].
class of preferences that are necessary and sufficient for this to happen (i.e. for the capital tax to be zero at all dates). This is then an intragenerational optimality condition.

The OLG economy thus allows the distinction between intergenerational optimality and intragenerational optimality, and this helps in the interpretation of the results.

We should notice that in a dynastic economy it is optimal to tax capital heavily at the beginning of the optimisation period (regardless individual heterogeneity). This is generally not the case in the overlapping generations economy as long as the old generation enters into the social welfare function.

The paper also analyses third best policy, but under the more realistic assumption that the government chooses fiscal policy that affect two generations at the same time, e.g. present wage tax and present capital income tax or next period’s wage tax and next period’s capital income tax. This is also new to the literature.

Table 1 gives a taxonomy of previous studies of optimal taxation in overlapping generations economies.

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Consumers</strong></td>
</tr>
<tr>
<td>Diamond (1973)</td>
</tr>
<tr>
<td>Ordover and Phelps (1979)</td>
</tr>
<tr>
<td>Atkinson and Sandmo (1980)</td>
</tr>
<tr>
<td>Park (1991)</td>
</tr>
<tr>
<td>This paper</td>
</tr>
</tbody>
</table>
The period-$t$ government controlling taxes that affect current generation only, e.g. the labour tax in period $t$ and the capital income tax in period $t+1$.

The period-$t$ government controls either the taxes at time $t$, or the taxes at time $t+1$.

There are some common characteristics of these studies.\(^2\)

(i) They are characterised by perfect competition and constant returns-to-scale in (aggregate) production. 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 (because future generations’ preferences are not reflected in current prices).\(^3\) However, the possibility of dynamic inefficiency only arises if there is population growth.

(ii) Consumers are heterogeneous, not only with respect to age, but also in abilities (Ordover and Phelps) and tastes (Diamond). In Atkinson and Sandmo individuals only differ with respect to age.

(iii) Consumers live for two periods and have no bequest motives.

(iv) There is physical non-perishable capital, which initial level is greater that zero.\(^4\)

(v) Each government can only set taxes that affect only one generation at each point in time. For example setting present labour taxes and future capital income taxes. This rules out the possibility of finding the second-best structure of taxes (which would require optimisation over all future taxes), and rules out the possibility of analysing governments’ decisions that affect two generations simultaneously.\(^5\)

Diamond (1973) analyses optimal taxation in a very general Overlapping-Generations (OLG) economy, allowing for many consumption goods (privately and publicly produced) in each period and labour supply also in the second period. Individuals may differ in tastes, there are no bequest motives, there are no pure public goods), there is a constant population growth

\(^2\) They are all, more or less, applications of the Diamond (1965) economy.

\(^3\) Cass (1972).

\(^4\) The Samuelson (1958) economy does not have capital, why individuals have no possibility in transferring consumption possibilities from the present to the future, other than (i) having a government doing so (transfer from young to old, pay-as-you-go pension system), or (ii) trading in fiat money.

\(^5\) Setting present labour taxes and present capital income taxes or future labour taxes and future capital income taxes would affect two generations simultaneously. This seems to be a more plausible formulation.
rate, and the government cares about all future generations, but can only tax the present generation.\textsuperscript{6}

The economy is studied under three regimes: (i) central planning (the government controls individuals’ quantities) which can be implemented through a decentralised economy with individual specific lump sum taxes, (ii) fully taxed economy (the government can tax all commodities), (iii) partially taxed economy (the government can tax only some commodities, or have to tax some commodities at the same rates).

However, Diamond’s focus is on the marginal products in public versus private production, rather than on the tax rules themselves. The main conclusions are that under all regimes, if the economy goes to a steady state, the marginal product of capital in public production equals the social discount rate, i.e. the Modified Golden Rule is obtained. Under regime (i) and (ii) this is also true for marginal product of capital in private production.

Ordover and Phelps (1979) use a simplified version of Diamond’s (1973) OLG economy.\textsuperscript{7} The difference is that in the latter framework there is only one consumption good available at each date, and that individuals work only in their first period of their lives, and that there is no public production. Ordover and Phelps also abstract from population growth. In their economy the government has access to non-linear labour income and capital income taxes, and the government can differentiate the lump-sum transfers between the young generation and the old.

Ordover and Phelps focus on the tax rules, and their main results are that the marginal tax on the highest income earned is zero (both wage and interest income), and if each worker’s utility is everywhere weakly separable between consumption in period one and two and in leisure, then the capital income tax is always zero.

Atkinson and Sandmo (1980) examine optimal linear taxes when individuals of the same generation are identical and the population grows at a constant rate. The basic framework is otherwise close to Ordover and Phelps. Individuals consume in two periods but work only in

\textsuperscript{6}The way in which Diamond solves the problem is by letting the present government controlling the prices affecting the present generation only. This implies that if for example the government has access to wage and interest taxes, then the present government chooses present wage tax and future interest tax.

\textsuperscript{7}They use the Diamond (1965) OLG economy which is the most common framework for tax analysis.
the first and the government chooses taxes that affect only one generation.

Among their results are: (i) if generational lump sum taxes are allowed the Modified Golden Rule holds in steady state and no distortionary taxes are used, (ii) if no lump-sum transfers are allowed and utility is additively separable and logarithmic, then if the steady state obeys the Modified Golden Rule the capital income tax is zero in that steady state, (iii) if on the other hand the modified golden rule cannot be reached, capital income may be taxed or subsidised in steady state.

Park (1991) uses the same framework of Atkinson and Sandmo but allows individuals to differ in preferences and in time endowments, and also performs the analysis with a consumption tax. Among the results are: (i) with identical logarithmic preferences (individuals differ only in time endowments) the optimum capital income tax is zero in steady state, (ii) with identical time endowments the optimum capital income tax is zero in steady state, (iii) if individuals differ in preferences and time endowments it may be positive, negative or zero.

To summarise: the above studies have found the optimal capital income tax to be zero in steady state if either: (i) individuals’ utilities are logarithmic in period-one and -two consumption and in leisure, or (ii) the government has a non-linear tax schedule and utilities are weakly separable between period-one and -two consumption and in leisure.

In light of these conclusions one may be tempted to think that it is a fundamental difference in optimal policy between OLG economies and dynastic economies, presumably because of the assumption of finitely lived households in OLG, (or equivalently absence of bequest motives) or because of the possibility of a dichotomy between the government’s and the individuals’ discount rates. However, as we shall see in this paper, this conclusion is not

---

8 Park intends to allow individuals to differ in abilities, using $h$ as notation for hours worked and $l$ as notation for effective labour supply. $l$ rightly enters the budget constrain, but $h$ does not enter the utility function, instead $l$ wrongly enters the utility. Thus, the paper does not allow for productivity differences. On the other hand Park allows individuals to differ in time endowments (i.e. total hours available for work and leisure) which is a completely different thing, since then individuals face the same wage rate.

9 The zero capital income tax has been derived in dynastic economies under less restrictive assumptions [Chamley (1986), Lucas (1990)].
correct. We will show that the optimality of the zero capital-income tax generality carries over to the OLG economy as well. The reason for the difference in the results between this paper and the above mentioned studies is not in the basic assumptions on the economic structure but in that the latter have not found the second-best optimal taxes, since the government is assumed to control the taxes affecting the present generation only. If we want to find the second best we have to optimise with respect to all future taxes. The results will in general differ because of the time-inconsistency of optimal policy, but the second best gives us a useful benchmark at which we can evaluate the time-consistent policy (usually referred to as "third best").

Since the second best (open loop) formulation generally gives a time-inconsistent solution, we would be interested in the recursive formulation, when the government cannot commit to the future sequence of taxes. However, in the above mentioned studies the recursive formulation is such that the present government chooses taxes that affect the present generation only, i.e. the present labour income tax and the next period’s capital income tax. It is more plausible that a government chooses policy instruments that affect both generations at the same time, e.g. present labour income tax and present capital income tax, or future labour income tax and future capital income tax. We will address this issue as well in this paper.

The paper is structured as follows:

In Section 2 a general OLG-economy is described. In Section 3 the optimal tax problem is solved when the government can choose all future taxes (full commitment). The government has access to a flat labour income tax and a flat capital income tax, and is able to pass on debt or assets to the next generation. It is proven that for social welfare being recursively separable across generations the optimal capital income tax is zero in steady state. We also derive the class of preferences which are necessary and sufficient for the capital income tax to be zero at all dates.

In Section 4 optimal taxation with no commitment is analysed. The modified Golden Rule result may break. A class of preferences are shown to guarantee the modified Golden Rule. We also show that when the utility function of the individuals are logarithmic, the sequence of optimal taxes in the third best coincides with the optimal tax sequence in the second best (i.e. the open-loop and closed-loop solutions coincide).

Section 5 deals with optimal taxation with one-period commitment. The steady state is
always characterised by the Modified Golden Rule. Furthermore, we derive the class of preferences that are necessary and sufficient to yield zero capital tax at all dates. They coincide with those giving the same result in the second best.

Section 6 shows that Theorem 1 is robust to a number of extensions. Section 7 concludes the paper.

2 THE ECONOMY

Individuals live for two periods. They consume both as young and as old, but work only when young. They have preferences over period-one consumption, period-one labour supply, period-two consumption, and (possibly) period-one and period-two provision of public goods.\(^\text{10}\) To account for population growth, let \(N_t\) be the size of the young generation at \(t\). Within each generation individuals differ (in productivity and possibly also in preferences), and the types, \(i\), are distributed according to \(F(i)\), which is continuous and stationary. This implies that while the population grows, the distribution within each generation is constant.\(^\text{11}\) We shall assume that the growth in \(N_t\) is constant and exogenous, \(N_t = (1+n)N_{t-1}\), and we normalise \(\Gamma\) such that \(\int d\Gamma(i) = 1\). A key difference between individuals within one generation is their productivity, which is represented by a skill parameter \(\gamma\).

In period one individual \(i\) born at \(t\) (with ability \(\gamma\)) supplies labour \(l_t^i\) on the market and consumes \(c_t^i\) units of the only consumption good. She is paid \(w_t\) per efficient unit of supplied labour, i.e. in proportion to \(\gamma l_t^i\), and she saves \(a_{t+1}^i\) for the next period. Let \(\tau_t^l\) and \(\tau_t^k\) denote the wage-income-tax rate and the capital-income-tax rate, respectively. In period two she receives the after-tax return, \(P_{t+1}\), on her savings which is used for consumption \(c_{t+1}^i\). The period-one and period-two provisions of the public good are denoted \(G_t\) and \(G_{t+1}\), respectively. It is convenient to structure the assumptions as follows:

\(^{10}\) Most of the analysis in this paper will abstract from bequests entirely. An extension to allow for bequests is considered in section 7.

\(^{11}\) See Samuelson (1956).
2.1 Assumptions

A1 Individual Preferences

The utility function

\[ U^H(t_i^H, l_t^H, c_{t+1}^H, g_t, g_{t-1}) \]  

is assumed to be strictly concave in all arguments, and \( g_t \) and \( g_{t+1} \) are per-capita public consumption.

A2 Individuals’ Constraints

The individual budget constraints are

\[ c_t^H + a_{t+1}^H = (1 - \tau^H) w_t l_t^H \]  

\[ c_{t+1}^H = [1 + (1 - \tau^H) r_{t-1}] a_t^H \]  

A3 Production

A large number of firms are operating under a constant-returns-to-scale technology. Therefore aggregate production, \( Y_t \), can be calculated as if there was a representative firm employing the aggregate quantities of capital and labour, defined as

\[ K_t = N_t \int k_t d\Gamma(i) \]  

\[ L_t = N_t \int l_t d\Gamma(i) \]  

respectively.  

\[ F(K, L) = F_K(K, L) K + F_L(K, L) L \]  

A4 Government’s Constraint

The government is allowed to borrow and lend freely at the market rate of interest and may take the aggregate expenditure requirement \( G_t, t=0,\ldots,\infty \), as \( (a) \) predetermined, \( (b) \) giving utility as a flow, \( (c) \) giving utility as a stock. Denote the aggregate government debt as \( B_t \).

\[ B_{t+1} = \left[ 1 + (1 - \tau^H) r_t \right] B_t - \tau^H r_t K_t - \tau^j w_t L_t + G_t \]  

12 Only the young generation work and only the old own capital.
A5 Government’s Objective
The government seeks to maximise a Bergson-Samuelson welfare function including all individuals of present and future generations.

A6 Population Growth
The population growth is exogenous (possibly zero)

\[ N_t = (1+n)N_{t-1} \]  

2.2 On the assumptions
A1. Consumption in both periods but only labour supply in the first intends to capture the idea that individuals eventually retire from work, and have to save as young for future consumption as old. This follows Diamond (1965), and is the most common way of modelling OLG economies. One of the key aspects is age heterogeneity, which adds a dimension to the optimal tax problem. Not only will taxes affect the individuals’ consumption allocation across two dates, but will also affect the allocation of consumption between generations.

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). The nature of these imperfections would be of crucial importance in determining an optimal tax structure. By incorporating imperfections there would be an additional motivation for taxation: corrective taxation. This paper focuses on taxes in a world where markets are perfect, and the taxes as causing inefficiency rather than correcting existing ones. See further the discussion of A3 below.

As mentioned above labour is supplied in the first period only. This "forces" positive savings (savings for retirement).
A3. Perfect competition, constant-returns-to-scale technology, and no production externalities. This is in order to focus on the efficiency-equity aspects of taxation, rather than an eventual corrective role. If there was imperfect competition, economies of scale, or production externalities, the tax rates would involve a "Pigouvian" element: corrective taxation.

A4. Most often in Ramsey-tax problems the government expenditure (in real terms) is exogenously determined. One 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. However, we would surely be interested in how the tax system and public expenditure interact, for example how public expenditure would be determined in high-debt economies, or in economies with restrictions on tax instruments etc. In order to do so we explicitly have to take into account the function of public expenditure. We will treat it as public consumption and we will focus on the Samuelson rule as a benchmark.  

A5. That the government cares about all future generations is in line with previous literature. The restrictions on the welfare function will be imposed later (in the relevant sections).

A6. We allow for a constant population growth to account for the most general case. It is well known that population growth may cause dynamic inefficiency, thus giving a role for corrective taxation. However we will not focus on this aspect.

In accordance with most of the previous literature we shall abstract from consumption taxation and assume that individuals have no preferences over leaving bequests.

We shall later on make more explicit assumptions than these and then make a deeper discussion of the simplifications.

13 We could have thought of public expenditure as entering the production technology. However, it would give us rather simple provision rules, not directly related to preferences. This extension is merely an exercise.
2.3 Economic Equilibrium

In this section it is described how individual and aggregate economic behaviour can be solved for, given any arbitrary sequences of tax rates and public expenditure. Define the gross return on capital as $R_t \equiv (1+r_t)$ and the after-tax prices as $P_t \equiv 1+(1-\tau_t^k)r_t$ and $\omega_t \equiv (1-\tau_t^l)w_t$. By profit maximisation the before-tax prices (the interest rate and the average wage rate) are given by $r=F_k$ and $w=F_L$.

**Individual Economic Behaviour**

Maximisation of (1) subject to (2)-(3) gives the individuals’ decision rules $\{c_{it}, l_{it}, c_{i,t+1}\}$ as functions of $\{\omega_t, P_{t+1}, g_t, g_{t+1}\}$ and indirect utilities

$$V^U = V^U(P_{t+1}, \omega_t, g_t, g_{t+1})$$ (7)

**Aggregate Economic Behaviour**

Because of the OLG structure of the economy it is not possible to have an aggregation theorem that gives a representative individual result.\(^{14}\) That is not a problem for the analysis, though.

Aggregate consumption (as well as all other quantities) may be rewritten in terms of per capita of the young generation. Define generation average consumptions as $\overline{c_t} \equiv \int c_{it} d\Gamma(i)$ and $\overline{c_{t+1}} \equiv \int c_{i,t+1} d\Gamma(i)$. Then

$$\overline{c_t} = \frac{C_t}{N_t} = \overline{c_t}(\omega_t, P_{t+1}, g_t, g_{t+1}) + \frac{\overline{c_{t+1}}(\omega_{t+1}, P_t, g_{t+1}, g_t)}{1+n}$$ (8)

Similarly for labour, let the average (in efficiency units) be $\overline{l_t} \equiv \int l_{it} d\Gamma(i)$. Then

---

\(^{14}\) When individuals have the same length of life, presumably infinite, and differ only in capital endowments, and productivity the aggregate economic behaviour coincides with a representative individual if the felicity function belongs to the additive or multiplicative HARA-family.
We, analogously, define per capita (old generation) debt as \( b_t \equiv B_t / N_t \), and per capita (old generation) capital as \( k_t \equiv K_t / N_t \), then, after some manipulation,

\[
\bar{l}_t = \frac{I_t}{N_t} - \bar{l}_t(\omega_t, P_{t+1}, g_t, g_{t+1})
\]

Finally, the aggregate capital stock evolves according to

\[
b_{t+1} = P_t \bar{k}_t \left( 1 + \frac{k_t}{1 + n} \right) - F \left( \frac{k_t}{1 + n}, \bar{l}_t \right) + \omega_t \bar{l}_t - \frac{k_t}{1 + n} + g_t
\]

Finally, the aggregate capital stock evolves according to

\[
k_{t+1} = \frac{k_t}{1 + n} + F \left( \frac{k_t}{1 + n}, \bar{l}_t \right) - \bar{c}_t - \frac{\bar{c}_{t+1}}{1 + n} - g_t
\]

Thus, the variables denoted by bars are average quantities within each generation and depend on the consumer prices, the amount of public goods and the distribution of bequests.

2.4 Steady State

Most of the paper considers the optimal taxes at a steady state. A steady state in an overlapping generations economy is taken to be the situation in which \( k_t = k, b_t = b, \tau^k = \tau^c, \tau^l = \tau, \forall t \). This implies that the consumption pattern of generation \( t \) is also the consumption pattern of generation \( t+1 \), but does not imply that consumption of an individual is constant over time. For example, each individual may have an increasing (decreasing) consumption path over her lifetime, still the economy would be at a steady state if the next generation have the same consumption path. This is an important difference from the dynastic models, where a steady state necessarily implies that consumption is constant over time. From an optimal-tax point of view, one might be tempted to think that a zero capital-income tax may be a result of the optimality of taxing consumption of different dates uniformly, if consumption is constant. As we will see later on this intuition is not true.
2.5 Welfare

In the previous literature the preferences of the social planner are utilitarian (sum of utilities) and thereby additively separable in generations and individuals. We shall keep our social welfare function more general than that, but give more structure than assumption A5. We shall assume

A5(a) The social welfare function is Bergson-Samuelson over individuals within each generation

\[ W_t = W(\{V^t_i\}) \]  

(12)

and of the recursively separable (Koopmans) form over generations

\[ J(W_t, W_{t+1}, W_{t+2},...) = U(W_t, J(W_{t+1}, W_{t+2},...)) \]  

(13)

Koopmans (1960) derived this functional form for a utility function for ordering consumption at different dates. We, instead, use the function for the ordering of social welfare of different generations. Koopmans derived the function from axioms on consumption orderings. Here the same axioms has to be translated into intergenerational welfare orderings. The key axioms are weak separability (limited independence) and time independende (stationary preferences). It is worthwhile to explore the key axioms here in the context of intergenerational welfare orderings. Let \( W = W_t \) (i.e. the Bergson-Samuelson welfare function of generation \( t \), as above). Further, let \( W = (W_{t+1}, W_{t+2},...) \) (i.e. the vector of welfare functions of future generations).

Next, compare two programmes \( W, W' \) and \( W, W' \).

**Axiom W**: Weak Separability ("limited independence")

\[ W, W \succeq W', W \iff W, W' \succeq W', W' \]

That is, the ordering of welfare between two different generations are independent of the welfare of the future generations.

**Axiom T**: Time Independence ("stationary preferences")

\[ W, W \succeq W', W \iff W \succeq W' \]

---

15 Recall that Chamley (1986) found the optimal capital-income tax to be zero in steady state for an economy where a representative individual has got a utility function of the Koopmans form.
That is, the ordering of future welfare alternatives are independent of the present welfare level. The latter implies that postponing decisions does not alter rank order (and thus ensures time-consistency in social preferences).

Finally, we need to say something about discounting future generations. The social discount rate, denoted $\delta_t$, is the marginal change in social welfare at date $t$ due to a change in welfare of generation $t+1$, divided by the marginal change in social welfare at date $t+1$ due to a change in welfare of generation $t+1$.

$$
\delta_{t+1} = \frac{\partial J(W_t, W_{t+1}, \ldots)}{\partial W_{t+1}} \left/ \frac{\partial J(W_t, W_{t+1}, \ldots)}{\partial W_{t+1}} \right. \frac{\partial U(W_t, J(W_{t+1}, W_{t+2}, \ldots))}{\partial J} \tag{14}
$$

If the ratio is smaller than unity, future generations’ welfare is discounted. This will be the assumption throughout the analysis.

The Koopmans form is desirable for the purpose of this paper. First, it contains a broad class of functions, that can be derived from axioms. The utilitarian is a special case (where $\delta$ is constant). Second, it also makes a steady state possible, and thereby steady-state analysis possible. Third, the social preferences are time-consistent. This ensures that if there is any difference between the open-loop and the closed-loop solutions, it is due to time-inconsistency problems resulting from the second-best framework, and not from the objective function itself. This makes the second- and third-best easily comparable. Fourth, weak separability and time-independence are not too unreasonable when ranking welfare of different generations.

It should be pointed out that all previous studies on optimal policy in overlapping-generations economies have assumed special cases of the Koopmans form (e.g. discounted sum of future generations’ utilities), therefore our treatment is more general than anyone else’s sofar.

We will now turn to the optimal tax problem. In finding the second-best optimal taxes we have to optimise over all future taxes, thus solving the open-loop formulation, assuming that the present government can "force" future governments to follow the optimal plan.
3 OPTIMAL INCOME TAXATION
- THE SECOND BEST

3.1 The optimal tax problem
We shall state the first result. The proof is instructive and is therefore included in the main text.

Theorem 1 Assume A1-A4, A5(a), A6, and that the government can enforce its optimal plan for the entire future. If the economy reaches a steady state, such that per capita quantities are constant, then the optimal capital income tax is zero in the steady state and the rate of interest satisfies the modified Golden Rule: $\delta R = 1 + n$.

Proof: We solve the open-loop problem (present government can "force" future governments to implement the present government’s optimal policy). The government chooses a sequence $\{\omega_t, P_t\}, t=0,...,\infty$, subject to the aggregate constraints (10) and (11) and subject to the individuals’ optimal behaviour. We may write the Lagrangean as

$$\mathcal{L} = J(W_{-1}, W_0, W_1, \ldots) + \sum_{t=0}^{\infty} \mu_t [B(\omega_t, P_t, P_{t+1}, b_t, k_t) - b_{t+1}]$$
$$+ \sum_{t=0}^{\infty} \lambda_t [K(\omega_{t-1}, \omega_t, P_t, P_{t+1}, k_t) - k_{t+1}]$$

(15)

where $B(\omega_t, P_t, P_{t+1}, b_t, k_t)$ and $K(\omega_{t-1}, \omega_t, P_t, P_{t+1}, k_t)$, are the right-hand-sides of (12) and (13) respectively. The first-order conditions are

$$\frac{\partial J}{\partial W_t} \frac{\partial W_t}{\partial \omega_t} + \mu_t \left[ I_t + (\omega_t - w_t) \frac{\partial I_t}{\partial \omega_t} \right] + \lambda_t \left[ w_t \frac{\partial I_t}{\partial \omega_t} - \frac{\partial \bar{c}_t}{\partial \omega_t} \right] - \frac{\lambda_{t-1}}{1 + n} \left[ \frac{\partial \bar{c}_{t-1}}{\partial \omega_t} \right] = 0$$

(16)
First we rewrite the derivatives of the welfare function $J()$. Differentiating (13) with respect to $W_t$, then after recursive substitution we have

$$
\frac{\partial J}{\partial W_t} \frac{\partial W_t}{\partial P_{t+1}} + \mu_t(\omega_t - w_t) \frac{\partial \tilde{l}_t}{\partial P_{t+1}} + \frac{\mu_{t+1}}{1+n}(b_{t+1} + k_{t+1}) + \lambda_t \left[ \frac{\partial \tilde{c}_t}{\partial P_{t+1}} - \frac{\partial \tilde{c}_t'}{\partial P_{t+1}} \right] - \frac{\lambda_{t+1}}{1+n} \left[ \frac{\partial \tilde{c}'_{t+1}}{\partial P_{t+1}} \right] = 0
$$

(17)

$$
- \mu_t + \frac{\mu_{t+1}}{1+n} P_{t+1} = 0
$$

(18)

$$
\frac{\mu_{t+1}}{1+n} (P_{t+1} - R_{t+1}) - \lambda_t + \frac{\lambda_{t+1}}{1+n} R_{t+1} = 0
$$

(19)

Furthermore, define $\tilde{\mu}_t \equiv \prod_{s=0}^{t} \delta_s l_{s+1}$, $\tilde{\lambda}_t \equiv \prod_{s=0}^{t} \delta_s \lambda_s$, then multiplying the first-order conditions (16)-(19) by $\prod_{s=0}^{t} \delta_s^{-1}$ we obtain\(^{16}\)

$$
\frac{\partial J(W_{t-1}, W_1, W_2, \ldots)}{\partial W_t} = \prod_{s=0}^{t} \frac{\partial U(W_{t-1}, J(W_s, W_{s+1}, \ldots))}{\partial J} \frac{\partial J(W_t, J(W_{t-1}, W_{t+1}, \ldots))}{\partial W_t}
$$

(20)

where the second equality follows from (14).

Furthermore, define $\tilde{\mu}_t \equiv \prod_{s=0}^{t} \delta_s^{-1} \mu_s$, $\tilde{\lambda}_t \equiv \prod_{s=0}^{t} \delta_s^{-1} \lambda_s$, then multiplying the first-order conditions (16)-(19) by $\prod_{s=0}^{t} \delta_s^{-1}$ we obtain\(^{16}\)

$$
\frac{\partial U}{\partial W_t} \frac{\partial W_t}{\partial \omega_t} + \tilde{\mu}_t \left[ \frac{\partial \tilde{l}_t}{\partial \omega_t} - (\omega_t - w_t) \frac{\partial \tilde{c}_t}{\partial \omega_t} \right] + \tilde{\lambda}_t \left[ \frac{\partial \tilde{c}'_t}{\partial \omega_t} + \lambda_t \frac{\partial \tilde{c}_t'}{\partial \omega_t} - \frac{\lambda_{t+1}}{1+n} \frac{\partial \tilde{c}'_{t+1}}{\partial \omega_t} \right] = 0
$$

(21)

\(^{16}\) Note that we have $\prod_{s=0}^{t} \delta_s^{-1} \mu_{t+1} = \delta_{t+1} \prod_{s=0}^{t} \delta_s^{-1} \mu_s = \delta_{t+1} \tilde{\mu}_{t+1}$.  

16
Substituting (23) and (24) into (21) and (22) gives, respectively,

\[
\frac{\partial U}{\partial W_t} \frac{\partial W_t}{\partial P_{t+1}} + \tilde{\mu}_t (\omega_t - w_t) \frac{\partial \tilde{I}_t}{\partial P_{t+1}} + \frac{\delta_{t+1} \tilde{\mu}_{t+1}}{1+n} (b_{t+1} + k_{t+1}) \\
+ \tilde{\lambda}_t \left[ w_t \frac{\partial \tilde{I}_t}{\partial P_{t+1}} - \frac{\partial \tilde{c}_t}{\partial P_{t+1}} \right] - \frac{\delta_{t+1} \tilde{\lambda}_{t+1}}{1+n} \left[ \frac{\partial \tilde{c}_t}{\partial P_{t+1}} \right] = 0
\]

\[
- \tilde{\mu}_t + \frac{\delta_{t+1} \tilde{\mu}_{t+1}}{1+n} p_{t+1} - 0
\]

\[
\frac{\delta_{t+1} \tilde{\mu}_{t+1}}{1+n} (p_{t+1} - r_{t+1}) - \tilde{\lambda}_t + \frac{\delta_{t+1} \tilde{\lambda}_{t+1}}{1+n} r_{t+1} = 0
\]

Substituting (23) and (24) into (21) and (22) gives, respectively,

\[
\frac{\partial U}{\partial W_t} \frac{\partial W_t}{\partial \omega_t} + \tilde{\mu}_t \left[ I_t + (\omega_t - w_t) \frac{\partial \tilde{I}_t}{\partial \omega_t} + \frac{p_{t+1} - r_{t+1}}{R_{t+1} p_{t+1}} \frac{\partial \tilde{c}_t}{\partial \omega_t} \right] + \tilde{\lambda}_t \left[ w_t \frac{\partial \tilde{I}_t}{\partial \omega_t} - \frac{\partial \tilde{c}_t}{\partial \omega_t} - \frac{1}{R_{t+1}} \frac{\partial \tilde{c}_t}{\partial \omega_t} \right] = 0
\]

\[
\frac{\partial U}{\partial W_t} \frac{\partial W_t}{\partial P_{t+1}} + \tilde{\mu}_t \left[ (\omega_t - w_t) \frac{\partial \tilde{I}_t}{\partial P_{t+1}} + \frac{p_{t+1} - r_{t+1}}{R_{t+1} p_{t+1}} \frac{\partial \tilde{c}_t}{\partial P_{t+1}} \right] \\
+ \tilde{\lambda}_t \left[ w_t \frac{\partial \tilde{I}_t}{\partial P_{t+1}} - \frac{\partial \tilde{c}_t}{\partial P_{t+1}} \right] = 0
\]

Differentiate the individuals’ intertemporal budget constraints, \( c_{t+1}^{\prime} + c_{t+1}^{\prime} / P_{t+1} = \omega_t l_t^\prime \), with respect to \( \omega_t \) and \( P_{t+1} \), and aggregate to obtain

\[
\frac{\partial \tilde{c}_t}{\partial \omega_t} - P_{t+1} \left[ \tilde{I}_t + \omega_t \frac{\partial \tilde{I}_t}{\partial \omega_t} - \frac{\partial \tilde{c}_t}{\partial \omega_t} \right]
\]
\[ \frac{\partial c_{t+1}}{\partial P_{t+1}} = a_{t+1} + P_{t+1} \left[ \omega_t \frac{\partial l_t}{\partial P_{t+1}} - \frac{\partial c_t}{\partial P_{t+1}} \right] \]  

(28)

Substituting (27) and (28) into (25) and (26), respectively, gives

\[ \frac{\partial U}{\partial W_t} \frac{\partial W_t}{\partial \omega_t} = (\bar{\mu}_t - \bar{\lambda}_t) \left[ \frac{w_t}{\partial \omega_t} \frac{\partial l_t}{\partial \omega_t} - \frac{\partial c_t^i}{\partial \omega_t} \right] - \frac{P_{t+1}}{R_{t+1}} \left( \bar{l}_t + \omega_t \frac{\partial l_t}{\partial \omega_t} - \frac{\partial c_t^i}{\partial \omega_t} \right) \]  

(29)

\[ \frac{\partial U}{\partial W_t} \frac{\partial W_t}{\partial P_{t+1}} = (\bar{\mu}_t - \bar{\lambda}_t) \left[ \frac{w_t}{\partial P_{t+1}} \frac{\partial l_t}{\partial P_{t+1}} - \frac{\partial c_t^i}{\partial P_{t+1}} \right] - \frac{a_{t+1}}{R_{t+1}} - \frac{P_{t+1}}{R_{t+1}} \left( \omega_t \frac{\partial l_t}{\partial P_{t+1}} - \frac{\partial c_t^i}{\partial P_{t+1}} \right) \]  

(30)

Since the individuals’ indirect utilities are increasing in the after-tax wage, the left-hand side of (29) is positive, and therefore $\bar{\lambda}_t - \bar{\mu}_t \neq 0$. Equation (29) also shows that $\bar{\lambda}_t - \bar{\mu}_t$ is constant at a steady state. Then (23) implies that at a steady state $1+n = \delta P$, which together with (24) gives $(\bar{\lambda}_t - \bar{\mu}_t)(R-P)/P = 0$, which can be fulfilled if and only if $P=R$ (since $\bar{\lambda}_t - \bar{\mu}_t \neq 0$). Finally (23) at $P=R$ gives the modified Golden Rule. QED

Under surprising generality we have derived the zero capital-income-tax result. Thus, we need not have infinitely lived individuals (or bequest motives) for this rule to be optimal, nor do we need separability in consumption and leisure, nor do we need intertemporal separability of the individual utility function, nor do we need non-linear tax schedules.

### 3.2 Interpretation

The reason for the zero-capital income tax result has little to do with optimality of uniform consumption taxes. Note also that the consumption of an individual need not be constant at the steady state. We can very well have (and most likely we have) $c_i^t \neq c_i^{t+1}$, at a steady state.

The reason for the result is, instead, that the government has to be indifferent along two margins of variation: (i) transferring of tax burden from one date to the other and (ii) transferring of physical resources from one date to the other. The intuition is directly linked to the proof of the Theorem.

(i) The multiplier $\bar{\mu}_t$ (on the constraint $b_t$) is the marginal social value of public debt at date $t$, and is usually negative since the government has to use distortionary taxes to fund
public expenditure. Since more public debt implies that more tax revenue has to be collected at the present date (everything else equal), $\bar{\mu}$ is the marginal excess burden at date $t$. The marginal social value of public debt at date $t+1$, evaluated at date $t$, is $\delta_{t+1}\bar{\mu}_{t+1}$, and is also equal to the marginal excess burden at time $t+1$ (evaluated at date $t$). Then the marginal rate of substitution of tax burden from date $t$ to date $t+1$ is: $\text{MRS}(b_t, b_{t+1}) = \delta_{t+1}\bar{\mu}_{t+1}/\bar{\mu}_t$. This is, thus, the marginal rate of substitution in transferring tax burden from one period to the other.

The marginal rate of transformation in transferring tax burden is the after-tax interest rate. An increase in public debt implies that it has to be paid back at a rate $1+r_{t+1}$, but since individuals pay a tax on their return on their holdings of public debt, $\tau_{t+1}$, the marginal rate of transformation in transferring tax burden from one period to the other is: $\text{MRT}(b_{t+1}, b_t) = 1+(1-\tau_{t+1})r_{t+1} = P_{t+1}$. At the optimum $\text{MRS}=\text{MRT}$, which is condition (23).

(ii) The multiplier $\bar{\lambda}_t$ is the marginal social value of physical resources (capital) at date $t$. The marginal social value of physical capital $t+1$, evaluated at date $t$, is $\bar{\lambda}_{t+1}$. The marginal rate at which capital can be transferred from one period to the other is the pre-tax interest rate. The marginal value of transferring capital from one period to the other is $\bar{\lambda}_t + \delta_{t+1}\bar{\lambda}_{t+1}R_{t+1}$ plus any contribution capital will make to tax revenue in the next period, $\tau_{t+1}r_{t+1}\delta_{t+1}(\bar{\mu}_{t+1})$. This is condition (24).

At a steady state the government must be indifferent in transferring tax burden as well as physical capital from one period to the other. This implies that their respective rate of transformation must be equal. This can only be the case when capital is untaxed.

To conclude, the zero-capital income tax result is entirely due to the fact that the government has to be indifferent transferring both capital and tax burden from one period to the other, and this can only happen when capital is untaxed. The result follows intergenerational optimality.

### 3.3 Zero Capital Tax at All Dates

Though the above result has nothing to do with uniform consumption taxes, we can have situations where uniform consumption taxes could be optimal, and then the capital tax should be zero at all dates (not just at a steady state). This, as in the static commodity tax literature, happens for a special class of preferences. We shall first characterise those preferences for the situation when all individuals of the same generation are the same (i.e. under no
intragenerational heterogeneity), and then the situation with intragenerational heterogeneity. Zero capital tax at all dates is due to intragenerational optimality (while the general case in section 3.1 is due to intergenerational optimality). The proof in the identical-individual case is included in the main text as it is instructive.

**Theorem 2** Assume A1-A4, A5(a), A6, and that the government can enforce its optimal plan for the entire future. When all individuals within a generation are the same, then necessary and sufficient for the optimal capital-income tax to be zero at all dates (except the first date) is that the individual utility function belongs to the following class:

(i) \( u(c_t, c_{t+1}, l) \) is homogenous in all its arguments,

(ii) \( u(\tilde{u}(c_t, c_{t+1}), l) \) where \( \tilde{u} \) is homogenous in \( c_t \) and \( c_{t+1} \).

**Proof:** The optimality conditions are as in the proof of Theorem 1. Those of interest are (29) and (30). When all individuals within a generation are the same, we can consider a representative individual, and we have

\[
\frac{\partial W_t}{\partial \omega_t} = \frac{\partial V^t}{\partial \omega_t} - \frac{\partial u^t}{\partial c_t} l_t \tag{31}
\]

\[
\frac{\partial W_t}{\partial P_{t+1}} = \frac{\partial V^t}{\partial P_{t+1}} - \frac{\partial u^t}{\partial c_t} a_{t+1} P_{t+1} \tag{32}
\]

Combining (29) and (30), and using (31) and (32), gives

\[
M_t(w_t - \omega_t) = \left[ \frac{R_{t-1} - P_{t+1}}{R_{t+1}} \right] \left[ P_{t+1} l_t \frac{\partial c_t^t}{\partial P_{t+1}} - a_{t+1} \frac{\partial c_t^t}{\partial \omega_t} - \omega_t M_t \right] \tag{33}
\]

where

\[
M_t = P_{t+1} l_t \frac{\partial l_t}{\partial P_{t+1}} - a_{t+1} \frac{\partial l_t}{\partial \omega_t} \tag{34}
\]

By differentiating (9) (see the appendix) (34) becomes
where subscript 1 and 2 refers to \( t \) and \( t+1 \) respectively, and \( D > 0 \) due to concavity of \( u \) (see the appendix for the definition of \( D \)). Necessary and sufficient for \( M_t=0 \) is that the expression within square brackets is zero. Define

\[
\Phi(c_1,c_2,l) = \ln\left(\frac{u_{c_2}}{u_{c_1}}\right)
\]

then the expression within square brackets can be written as

\[
\Phi(c_1,c_2,l) c_1 + \Phi(c_2,c_1,c_2,l) c_2 + \Phi(c_1,c_2,l) l = 0
\]

Thus, whenever the partial derivatives are non-zero, \( \Phi \) must be homogenous of degree zero, i.e. \( \Phi(hc_1, hc_2, hl) = \Phi(c_1, c_2, l) \), for any constant \( h>0 \). Implies that \( uc_2(c_1,c_2,l)/uc_1(c_1,c_2,l) \) is homogenous of degree zero, i.e. \( uc_2(c_1,c_2,l) \) is homogenous of the same degree as \( uc_1(c_1,c_2,l) \), implying that \( u(c_1,c_2,l) \) is homogenous of any degree. This gives (i).

When \( \Phi_f(c_1,c_2,l) = 0 \), we must have \( uc_2(c_1,c_2,l)/uc_1(c_1,c_2,l) \) independent of \( l \), which only happens if \( u \) is weakly separable in \( l \), i.e. \( u=u(\tilde{u}(c_1,c_2),l) \). Then \( \Phi(hc_1, hc_2) = \Phi(c_1, c_2) \) for any constant \( h>0 \) implies that \( \tilde{u}c_2(c_1,c_2)/\tilde{u}c_1(c_1,c_2) \) is homogenous of degree zero, in turn implying that \( \tilde{u}(c_1,c_2) \) is homogenous of any degree. This gives (ii).

When \( \Phi_{c_2}(c_1,c_2,l) = 0 \), we must have the marginal rate of substitution between \( c_1 \) and \( c_2 \) is independent of \( c_2 \), which happens only if there is no income effect on \( c_1 \), implying that all income effect must be on \( c_2 \), and therefore utility must be linear in \( c_2 \); i.e. \( u=u(\tilde{u}(c_1,l) + c_2) \). Then \( \Phi(c_1,l) = 1/\tilde{u}c_1(c_1,l) \) is homogenous of degree zero, implying that \( \tilde{u}(c_1,l) \) is homogenous of degree one. However, since this implies that \( u \) is homogenous of degree one, it is a special case of (i).

When \( \Phi_{c_1}(c_1,c_2,l) = 0 \), we must have the marginal rate of substitution between \( c_1 \) and \( c_2 \) independent of \( c_1 \), and therefore utility must be linear in \( c_1 \); i.e. \( u=u(c_1 + \tilde{u}(c_2,l)) \). Then \( \Phi(c_2,l) = \tilde{u}c_2(c_2,l) \) is homogenous of degree zero, implying that \( \tilde{u}(c_2,l) \) is homogenous of degree one. Again, \( u \) is homogenous of degree one, and consequently a special case of (i). QED
As we see from (35), one requires that the various elasticities of the utility function cancel. Exactly this condition has been derived in the context of optimal commodity taxation with a representative consumer, Atkinson and Stiglitz (1972). When that condition is fulfilled uniform commodity taxation is optimal. In fact Atkinson and Stiglitz draw the conclusion, that when the various commodities are interpreted as dated goods, one should not tax capital. The class of preferences implying this was not derived, however.

Theorem 2 is a result concerning intragenerational optimality. Each consumer of each generation should face a uniform consumption tax, i.e. a zero tax on capital income.

We will now consider the case of heterogeneous individuals. In the commodity tax literature, it may happen that a many-person Ramsey rule reduces to a single-person Ramsey rule. In addition to the restrictions in Theorem 2, we would need restrictions such that the tax rules in a many-person economy reduces to those of a single person economy.

**Theorem 3** Assume A1-A4, A5(a), A6, and that the government can enforce its optimal plan for the entire future. When individuals within a generation differ in labour productivity only, then necessary and sufficient for the optimal capital-income tax to be zero at all dates (except the first date) is that the individual utility function belongs to the following class:

(i) \( u(c_t, c_{t+1}, l) \) is homogenous in all its arguments,

(ii) \( u(\bar{u}(c_t, c_{t+1}), l) \) where \( \bar{u} \) is homogenous in \( c_t \) and \( c_{t+1} \).

**Proof:** See the appendix.

It turns out that, to guarantee the zero-capital income tax result at all dates, individuals need to supply the same amount of \( \gamma \frac{l_t}{a_{t+1}} \), that is, the ratio of labour supply in efficiency units to the supply of savings must be invariant with respect to the underlying heterogeneity. It turns out that the same restrictions on \( u \) in Theorem 2 guarantees this. When this ratio is the same for all individuals, savings contain no extra information regarding \( \gamma \) than labour income \((w(\gamma l_t))\) does. This implies that a tax on savings should not be different from the tax that should be levied when all individuals have the same \( \gamma \), i.e. a zero tax on savings.
Finally, the class of utility functions resulting in Theorem 2 and 3, may look restrictive. We should notice, however, that many utility functions used in applied and computational work satisfy fall into this class (e.g. additively separable in all arguments and iso-elastic or logarithmic in consumption). The utility function logarithmic in consumption and leisure in work by Atkinson and Sandmo (1980, and Park (1981), clearly belongs to this class.

4 OPTIMAL INCOME TAXATION - NO COMMITMENT

In solving for the second best tax problem we assumed that the governemt could enforce its future tax sequence. When the government cannot commit itself (or committing future governments) to follow the second best plan, the government (or the future government) would generally find it optimal to deviate from the second best plan, and re-optimise. Then the government today would have to take that into account, when choosing policy today. Solving the optimal tax problem period by period would give us the time-consistent optimal taxes, i.e. the third-best optimum. The distinction between second best and third best is not due to changes in social preferences over time (the recursively separable social welfare function does not cause any time-inconsistency problems in itself). It is rather due to the fact that elasticities of tax bases are dependent on when the decision about a particular tax is taken.

4.1 Preliminaries

We shall now solve the problem for a social planner who cannot precommit. At each point in time the optimal tax policy is a function of the current state variables. The government optimal policy is a feed-back strategy. The present government takes the future feed-back strategies as given, but realises that it can influence future decisions by affecting the state variables. The relevant state for a time-$t$ government is the capital stock at $t$ and government debt at $t$. Rather than working directly with those states, we will transform one of the state variables.

Because of consumer heterogeneity, each government would have a vector of assets as a state, each element being the assets of a particular individual. We can reduce the number of states in the following way. We first replace the aggregate capital stock with the after-tax wage in the previous period. At date $t$ the capital stock is the old generation’s savings minus public debt. The old generation’s savings (assets) is a function of past after-tax wage $\omega_{t-1}$, and
expected future after-tax return on savings \( P_t^e \). That is

\[ a_t = k_t + b_t = a(\omega_{t-1}, P_t^e). \]

Taking the state variables at date \( t \) as \((\omega_{t-1}, b_t)\). Government policy at \( t \) is a pair \((\omega_t, P_t)\), which is a feed-back strategy \( \mathcal{E} \) from the state, i.e.

\[ (\omega_t, P_t) = \mathcal{E}(\omega_{t-1}, b_t). \]

Individuals of generation \( t-1 \) rationally predict \( P_t \) as a policy function \( P_t^e = P^e(\omega_{t-1}, b_t) \). In that way we may write the capital stock as

\[ k_t = a(\omega_{t-1}, P^e(\omega_{t-1}, b_t)) - b_t, \]

and we can replace \( k_t \) everywhere. Since each individual’s savings is a function of \( \omega_{t-1} \) and \( P^e(\omega_{t-1}, b_t) \), we need not to enter each individual’s assets as a separate state.

Next, replacing \( k_t \) everywhere in (10) we have \( b_{t+1} \) as a function of controls \((\omega_t, P_t)\) and states \((\omega_{t-1}, b_t)\), only:

\[
 b_{t+1} = \frac{(P_t - 1)a(\omega_{t-1}, b_t) + b_t}{1 + n} - F \left( \frac{a(\omega_{t-1}, b_t) - b_t}{1 + n}, \bar{l}(\omega_t, P_t^e) \right) + \omega_t \bar{l}(\omega_t, P_t^e) + g_t, \tag{38}
\]

where \( a(\omega_{t-1}, b_t) \) is shorthand for \( a(\omega_{t-1}, P^e(\omega_{t-1}, b_t)) \), and \( P_t^e = P^e(\omega_t, b_{t+1}) \). Denote

\[
 d_t = 1 - (\omega_t - w_t) \frac{\partial \bar{l}_t}{\partial P_t^e} \frac{\partial P_t^e}{\partial b_{t+1}}, \tag{39}
\]

Differentiating with respect to the policy variables we obtain

\[
 d_t \frac{\partial b_{t+1}}{\partial \omega_t} = \bar{l}_t + (\omega_t - w_t) \left( \frac{\partial \bar{l}_t}{\partial \omega_t} + \frac{\partial \bar{l}_t}{\partial P_t^e} \frac{\partial P_t^e}{\partial \omega_t} \right), \tag{40}
\]

\[
 d_t \frac{\partial b_{t+1}}{\partial P_t} = \frac{a_t}{1 + n}. \tag{41}
\]

Differentiating with respect to the states

\[
 d_t \frac{\partial b_{t+1}}{\partial b_t} = \frac{P_t - R_t}{1 + n} \frac{\partial a_t}{\partial P_t^e} \frac{\partial P_t^e}{\partial b_t} + \frac{R_t}{1 + n}, \tag{42}
\]

24
We also need the individuals’ indirect utilities. For an individual born at $t-1$, the period-$t$ government faces $V^t_{it-1} = u^{t-1}(ct_{it-1}, \omega_{it-1}, P_t, a_t^i, \gamma_{it-1})$, where $P_t(\omega_{it-1})$. At time $t$, $\omega_{it-1}, P_t$ are fixed and taken as given. This implies $\frac{\partial V^t_{it-1}}{\partial P_t} = (\frac{\partial u^{t-1}}{\partial c_{it-1}})a_t^i = (\frac{\partial u^{t-1}}{\partial \omega_{it-1}})\frac{\partial a_t^i}{\partial P_t}$. Furthermore, $\frac{\partial V^t_{it-1}}{\partial P_t} = 0$, and thereby $\frac{\partial V^t_{it-1}}{\partial \omega_{it-1}} = 0$.

Finally, $\frac{\partial V^t_{it-1}}{\partial \omega_{it-1}} = (\frac{\partial u^{t-1}}{\partial c_{it-1}})(\frac{\partial c_{it-1}}{\partial \omega_{it-1}}) + (\frac{\partial u^{t-1}}{\partial \omega_{it-1}})\frac{\partial c_{it-1}}{\partial \omega_{it-1}} + (\frac{\partial u^{t-1}}{\partial \omega_{it-1}})\frac{\partial a_t^i}{\partial \omega_{it-1}} = (P_t - \gamma_{it-1})(\frac{\partial u^{t-1}}{\partial c_{it-1}})(\frac{\partial c_{it-1}}{\partial \omega_{it-1}}) + (P_t - \gamma_{it-1})(\frac{\partial u^{t-1}}{\partial \omega_{it-1}})$.

Notice that the first term cancel at $P_t = \gamma_{it-1}$. We may then write the instantaneous welfare function as $W_{it-1} = W_{it-1}(P_t, \omega_{it-1}, b_t)$, the derivatives of which are

$$\frac{\partial W_{it-1}}{\partial P_t} = \int q_{t-1}^i a_t^i d\Gamma(i), \quad \frac{\partial W_{it-1}}{\partial b_t} = 0, \quad \frac{\partial W_{it-1}}{\partial \omega_{it-1}} = \int q_{t-1}^i \gamma_{it-1} d\Gamma(i) \quad (44)$$

where

$$q_{t-1}^i = \frac{\partial W_{it-1}}{\partial u^{t-1}} \frac{\partial u^{t-1}}{\partial c_{it-1}} P_t \quad (45)$$

### 4.2 The Optimal Tax Problem

We write the recursive problem as

$$J(\omega_{t-1}, b_t) = \max_{\omega_t, P_t} U(W_{t-1}(\omega_{t-1}, P_t), J(\omega_t, b_{t-1})) \quad (46)$$

s.t. $b_{t+1} = B(\omega_t, P_t, \omega_{t-1}, b_t)$

where the derivatives of $B(\cdot)$ are given by (40)-(43).

The first-order conditions are:

$$J_{\omega_t} + J_{b_{t+1}} \frac{\partial b_{t+1}}{\partial \omega_t} = 0 \quad (47)$$

25
Theorem 4 Assume A1-A4, A5(a), A6, and that the government cannot commit to future taxes. If the economy reaches a steady state and if either

(i) the individuals’ utility functions are: \( u = \ln(c_t-m(l_t)) + \beta \ln c_{t+1}, \) or \( u = \ln(c_t) + m(L-l_t) + \beta \ln c_{t+1}, \) or \( u = \ln(c_t) - \eta l_t + \beta \ln c_{t+1}, \) or

(ii) individuals’ expectation about future after-tax return on capital are independent of the level of public debt, or both,

then the steady state is characterised by the modified Golden Rule. Otherwise not necessarily.

Proof: Follows by substituting (39) and (42) into (50). If the utility function is of the form stated in the theorem \( \frac{\partial l_t}{\partial P_t} = \frac{\partial \omega_t}{\partial P_t} = 0, \) and \( \delta R = 1+n \) at the steady state (since \( U_2 = \delta \)). Part (ii) follows in the same way. QED

Theorem 5 Assume A1-A4, A5(a), A6, and that the government cannot commit to future taxes, then the optimal capital income tax is zero at all dates if the individuals’ utility function is either \( u = \ln(c_t) + m(L-l_t) + \beta \ln c_{t+1}, \) or \( u = \ln(c_t) - \eta l_t + \beta \ln c_{t+1}. \) The labour tax is identical the same as in the second best (i.e. under full precommitment).

Proof: See the appendix.

Theorem 5 implies that the closed-loop (feedback) solution coincides with the open-loop (control) solution when the individuals’ utility functions are logarithmic. In both cases the capital income tax is zero at all dates and the after-tax wage follows the first-order difference equation: \( \omega_{t+1} = \omega \delta R/(1+n). \)
5 OPTIMAL INCOME TAXATION - PARTIAL COMMITMENT

We will now turn to the case when the government can precommit to respective taxes one period in advance. Policy at $t+1$ is committed to at $t$. Therefore at $t$, $\omega_{t+1}$, $\omega_t$, and $P_t$ are given. The state equation for public debt is still given by (38). We treat $\omega_{t+1}$ as a state distinct from $\omega_t$. The controls of the government are $\omega_{t+1}$ and $P_{t+1}$.

We write the recursive problem as

$$J(\omega_{t-1}, \omega_t, P_t, b_t) = \max_{\omega_{t-1}, P_{t-1}} U(W_{t-1}(\omega_{t-1}, P_t), J(\omega_{t}, \omega_{t+1}, P_{t+1}, b_{t+1}))$$

s.t. $b_{t+1} = B(\omega_t, P_t, \omega_{t+1}, b_t)$

The necessary conditions (using $U_2 = \delta_t$) are

$$\delta_t \frac{\partial J_{t-1}}{\partial \omega_{t-1}} = 0$$

$$\delta_t \frac{\partial J_{t-1}}{\partial P_{t-1}} + \delta_t \frac{\partial J_{t-1}}{\partial b_{t-1}} \frac{\partial b_{t-1}}{\partial P_{t-1}} = 0$$

$$\frac{\partial J_t}{\partial \omega_t} - \delta_t \frac{\partial J_{t-1}}{\partial b_{t-1}} \frac{\partial b_{t-1}}{\partial \omega_t} + \delta_t \frac{\partial J_{t-1}}{\partial \omega_t}$$

$$\frac{\partial J_t}{\partial P_t} - U_1 \frac{\partial W_{t-1}}{\partial P_t} + \delta_t \frac{\partial J_{t-1}}{\partial b_{t-1}} \frac{\partial b_{t-1}}{\partial P_t}$$

$$\frac{\partial J_t}{\partial b_t} - \delta_t \frac{\partial J_{t-1}}{\partial b_{t-1}} \frac{\partial b_{t-1}}{\partial b_t}$$

$$\frac{\partial J_t}{\partial \omega_{t-1}} - U_1 \frac{\partial W_{t-1}}{\partial \omega_{t-1}} + \delta_t \frac{\partial J_{t-1}}{\partial b_{t-1}} \frac{\partial b_{t-1}}{\partial \omega_{t-1}}$$

**Theorem 6** Assume A1-A4, A5(a), A6, and that the government takes present taxes as given and can only commit to taxes one period ahead. If the economy reaches a steady state then the steady state is characterised by the modified Golden Rule. The capital-income tax is not necessarily zero.
Proof: Substituting for the derivatives of (38) into (53) and (54) we obtain

\[ \frac{\partial J_{t+1}}{\partial P_{t+1}} + \frac{\partial J_{t+1}}{\partial b_{t+1}} (\omega_t - w_t) \frac{\partial \bar{I}_t}{\partial P_{t+1}} = 0 \]  

(58)

\[ \frac{\partial J_{t+1}}{\partial b_{t+1}} \left[ \bar{I}_t + (\omega_t - w_t) \frac{\partial \bar{I}_t}{\partial \omega_t} \right] + \frac{\partial J_{t+1}}{\partial \omega_t} = 0 \]  

(59)

Using (58) in (55) (and the derivatives of (38)) we obtain

\[ \frac{\partial J_t}{\partial P_t} = U_t \frac{\partial W_{t+1}}{\partial P_t} - \delta \frac{\partial J_{t+1}}{\partial b_{t+1}} \left[ \frac{a_t}{1+n} + \frac{P_t - R_t}{1+n} \frac{\partial a_t}{\partial P_t} \right] \]  

(60)

Similarly for (56)

\[ \frac{\partial J_t}{\partial b_t} = \frac{\partial J_{t+1}}{\partial b_{t+1}} \frac{\delta R_t}{1+n} \]  

(61)

The modified Golden Rule follows from (61) at a steady state. QED

It is thus enough with one period precommitment in order for the steady state to satisfy the modified Golden Rule. As in the second-best case, the capital tax may happen to be zero at all dates:

Theorem 7 Assume A1-A4, A5(a), A6, and that the government takes the present taxes as given and can only commit to taxes one period ahead. When individuals within a generation are the same, or differ in labour productivity only, then necessary and sufficient for the optimal capital-income tax to be zero at all dates (except the first date) is that the individual utility function belongs to the following class:

(i) \( u(c_t, c_{t+1}, l) \) is homogenous in all its arguments,

(ii) \( u(\bar{u}(c_t, c_{t+1}), l) \) where \( \bar{u} \) is homogenous in \( c_t \) and \( c_{t+1} \).

Proof: Use the derivative of (38) in (57)
Substituting (61) into (60) and (62) and shifting forward we have

\[
\frac{\partial J_{t+1}}{\partial P_{t+1}} = U_i(t+1) \frac{\partial W_i}{\partial P_{t+1}} + \frac{\partial J_{t+1}}{\partial b_{t+1}} \left[ \frac{a_{t+1}}{R_{t+1}} + \frac{P_{t+1} - R_{t+1}}{R_{t+1}} \frac{\partial a_{t+1}}{\partial P_{t+1}} \right] \tag{63}
\]

\[
\frac{\partial J_{t+1}}{\partial \omega_t} = U_i(t+1) \frac{\partial W_i}{\partial \omega_t} + \frac{\partial J_{t+1}}{\partial a_{t+1}} \left[ \frac{P_{t+1} - R_{t+1}}{R_{t+1}} \frac{\partial a_{t+1}}{\partial \omega_t} \right] \tag{64}
\]

Using (63) and (64) in (58) we have

\[
\frac{\partial W_i}{\partial P_{t+1}} = -\frac{1}{U_i(t+1)} \frac{\partial J_{t+1}}{\partial b_{t+1}} \left[ \frac{a_{t+1}}{R_{t+1}} + \frac{\partial l_{t+1}}{\partial P_{t+1}} + \frac{P_{t+1} - R_{t+1}}{R_{t+1}} \frac{\partial a_{t+1}}{\partial P_{t+1}} \right] \tag{65}
\]

Using (63) and (64) in (59)

\[
\frac{\partial W_i}{\partial \omega_t} = -\frac{1}{U_i(t+1)} \frac{\partial J_{t+1}}{\partial b_{t+1}} \left[ \frac{\partial l_{t+1}}{\partial \omega_t} + \frac{\partial t_{t+1}}{\partial \omega_t} + \frac{P_{t+1} - R_{t+1}}{R_{t+1}} \frac{\partial a_{t+1}}{\partial \omega_t} \right] = 0 \tag{66}
\]

The terms within the square brackets in (65) and (66) are the negative of the terms within the square brackets in (29) and (30), respectively. Since \( U_i = \partial U / \partial W_i \), the rest of the proof is the proofs of Theorem 2 and 3. \( \text{QED} \)

7 EXTENSIONS

Bequests and Consumption Taxation

We can extend the analysis to bequests and consumption taxation in the following way.

An individual receives a bequest \( m_{it} \) when young, and pays a consumption tax, \( \tau_c \), proportional to consumption, and leaves a bequest \( m_{it+1} \). The utility function is now

\[
U^m(c_t, l_t, c_{t+1}, m_{t+1}, g_t, g_{t+1}, \ldots) \tag{67}
\]

and the individual’s budget constraints are
The key feature concerning the bequest motive is that the bequests themselves enter into the parents’ utility functions, not the children’s utilities. This breaks the links between marginal rates of substitution between different commodities at dates sufficiently distant apart (more than one period apart). This way of modelling bequests is not novel to the literature and provides a way of linking generations together in a non-dynastic way. Since the parents do not care about future generations’ utilities when choosing own consumption and labour supply, this framework may generate dynamic inefficiency if there is population growth (see comment on A6 in section 2).

Maximisation of (67) subject to (68)-(69) gives the individuals’ decision rules \{c^i_t, l^i_t, c^i_{t+1}, m^i_{t+1}\} as functions of \{ω_t, τ^c_t, τ^c_{t+1}, P_{t+1}, m^i_t, g_t, g_{t+1}\} and indirect utilities

\[ V^{it} = V^{it}\{P_{t+1}, ω_t, m^i_{t+1}, \tau^c_t, \tau^c_{t+1}, g_t, g_{t+1}\} \]

Bequests consists of the aggregate amount given by the old generation. Define the average bequests as \( \bar{m}_{t+1} \equiv \int m^i_{t+1} dF(i) \), then the aggregate bequests

\[ M_{t+1} = N_t \bar{m}_{t+1} = N_t \bar{m}_{t+1}\{P_{t+1}, ω_t, τ^c_t, τ^c_{t+1}, g_t, g_{t+1}, \{m^i_t\}\} \]  

depend on the distribution of \( m^i \) in the previous period. Therefore, in general, the aggregate bequests is not a meaningful concept. It is only if we make assumptions such that the aggregate is independent of the distribution equation (71) is useful. Next, aggregate consumption is

\[ \bar{c}_t = \frac{C_t}{N_t} = \bar{c}_t\{P_{t+1}, ω_t, τ^c_t, τ^c_{t+1}, g_t, g_{t+1}, \{m^i_t\}\} + \frac{\bar{c}_{t-1}\{P_t, τ^c_t, g_t, \{m^i_{t-1}\}\}}{1+n} \]

and similarly labour (in efficiency units)

\[^{17} \text{Saint-Paul and Verdier and others.} \]
Finally, the government’s budget constraint is

$$\bar{I}_t = \frac{L_t}{N_t} - \bar{I}(\bar{P}_{t-1}, \omega_t, \bar{\tau}_t, \bar{\tau}_{t-1}, g_t, g_{t-1}, \{m^u_t\})$$

(73)

where aggregate consumption is $C_t \equiv N_t \int c^t_i d\Gamma(i) + N_{t-1} \int c^t_{i-1} d\Gamma(i)$.

Suppose that individuals’ characteristics are perfectly inheritable. Perfect inheritability implies that a child with characteristic $i$ will inherit $m^u_t$, which has been bequeathed by a person with the same characteristic $i$. This implies that we may write the law of motion for bequests as a set of first-order difference equations

$$m^u_{t+1} = m^u_t \left\{ P_{t+1}, \omega_t, m_t^u, \bar{\tau}_t, \bar{\tau}_{t-1}, g_t, g_{t-1} \right\}, \forall i$$

(75)

Associate $\psi^i$ with each constraint (75), then the extra term in the Lagrangean (15) is

$$\sum_{i=0}^{\infty} \int \psi^i_t \left[ m^u_t \left\{ P_{t-1}, \omega_t, m_t^u, \bar{\tau}_t, \bar{\tau}_{t-1}, g_t, g_{t-1} \right\} - m^u_{t+1} \right] dF(i)$$

(76)

The extra term (76) does not change the key steps in the proof of Theorem 1, therefore the zero capital income tax still holds in steady state even if monetary bequests are included. It is obvious that it is true also when the consumption tax is included. The assumption underlying (75) and thereby also (76) is perfect inheritability. If this was not true and we would have randomization of characteristics, the specification of the problem becomes much harder. Essentially we would not have a steady state of the kind we assumed in the proofs of the Theorems. A way around the specification problem is to assume preferences such that, whatever randomization, the aggregate bequests are independent on the distribution, so that we would have [from equation (75)]

$$M_{t+1} = N_t \bar{m}_{t+1} = N_t \bar{m}_{t+1} \left\{ P_{t+1}, \omega_t, \bar{\tau}_t, \bar{\tau}_{t-1}, g_t, g_{t-1}, \bar{m}_t \right\}$$

(77)

and that the aggregate behaviour of consumption and labour supply (and thereby savings) is dependent only on aggregate bequests. Then, again, the steps in the proof of Theorem 1 remain unchanged and the optimum capital income tax is zero in the steady state. When we
include monetary bequest motives in the endogenous tax problem we will rely on the second route, i.e. when the aggregate is only a function of the aggregate, so we need not assume that the individuals’ characteristics are perfectly inheritable to render the problem tractable.

7 SUMMARY AND CONCLUSIONS

In this paper we have solved for the optimal tax rules in an overlapping generations economy, assuming different degrees of commitment: a) full commitment (second best), b) no commitment, c) partial commitment (next periods’ taxes are chosen today).\(^\text{18}\)

It was shown that the second best optimal capital income tax is zero in steady state if the social welfare function is recursively separable across generations. That steady state also obeys the Modified Golden Rule. Moreover we derived the class of preferences which are necessary and sufficient for the capital income tax to be zero at all dates.

In case b) the optimal capital income tax was generally not zero, and the steady state was generally not the Modified Golden Rule. The Modified Golden Rule was obtained as a special case if either (i) the individual’s utility functions belong to a restrictive class (logarithmic) or (ii) the level of public debt does not affect the individuals’ expectations about the after-tax return on capital (i.e. if the government’s feedback rule for the after-tax price on savings is independent of the level of public debt).

In case c) the capital income tax was generally not zero, but the steady state was always the Modified Golden Rule. However, exactly the same class of preferences that give zero-capital income tax at all dates in the second best, also give zero capital income tax at all dates in the third best.

---

\(^{18}\) No other study has previously solved cases a), b) or c), but instead a hybrid between b) and c): the present government chooses present wage tax and next period’s capital income tax. These studies have found the zero capital income tax only in steady state in the logarithmic utility case.
REFERENCES


