Taxing capital income in a perpetual youth economy

Valeria De Bonis[1]
Dipartimento di Scienze Economiche, Università di Pisa, Italy
debonis@ec.unipi.it

Luca Spataro[2]
Dipartimento di Scienze Economiche, Università di Pisa, Italy
l.spataro@ec.unipi.it

24 July 2003
Preliminary version

[1]Corresponding author: Dipartimento di Scienze Economiche, via C. Ridolfi 10, 56124, Pisa, Italy. Tel.: +39 050 2216382, fax: +39 050 598040.
[2]Tel.: +39 050 2216333, fax: +39 050 598040.

Abstract

We reconsider the issue of capital income taxation in an OLG perpetual youth framework. We show that the long run zero tax result does not generally hold. Besides the "life-cycle" motive pointed at by recent works, this work unveils two other forces pushing toward the taxation of capital income: the disconnection between generations and the relationship between the government and the individual intertemporal discount rates. We also show as a special case, that the non zero tax result applies also if age dependent taxes are not available, provided that the life-cycle behavior is ruled out, which cannot happen in the standard OLG models. Finally, it emerges that unfair life insurance contracts do not qualitatively affect the results.

Journal of Economic Literature Classification Numbers: E62, H21.

Key words: optimal dynamic taxation, primal approach, perpetual youth.

1 Introduction

Since the seminal works by Judd [14] and Chamley [6], there has been a growing number of contributions dealing with the issue of dynamic optimal capital income taxation. In particular, these two authors argued that the long run tax rate on capital income should be zero. This somehow striking result has been clarified only recently by a few works that have, on the one hand, highlighted the strict similarity with the more traditional static optimal taxation principles and, on the other hand, formally derived the conditions under which it can hold. In particular, Judd [15] has shown that the zero tax rate result descends directly from the fact that a tax on capital income is equivalent to a tax on future consumption increasing over time: thus, capital income should not be taxed if the elasticity of consumption is constant over time. However, as far as infinitely lived representative agent (ILRA) models are concerned[1], while this condition is necessarily true in the steady state, along the transition path, instead, it holds only if the utility function is (weakly) separable in consumption and leisure and homothetic in consumption. Moreover, both De Bonis and Spataro [9] and Erosa and Gervais [10][2] point out that, when separability is assumed out, the violation of the zero tax principle stems from the well known Corlett-Hague [8] rule: since leisure cannot be taxed directly, the second best solution is to tax (subsidize) the good that is more (less) complementary to it, i.e. consumption.

A further insight into the mechanism driving the mentioned result has been given by the adoption of the Overlapping Generation models with life cycle (OLG-LC). As shown by a number of authors[3], in this setup a non zero tax rate result holds in general, even in the long run, since the optimal consumption and labor plan is not generally constant over life, because of life-cycle behavior[4].

[1] See Atkeson et al. [1]. Among other articles focusing on the optimal capital income taxation problem see Jones et al. [13], dealing with human capital accumulation, and Chari et al. [7], Zhu [22] and Yakadina [21], dealing with stochastic frameworks.

[2] Both articles adopt the primal approach to the Ramsey problem; however, the former deals with an ILRA model, while the latter with an overlapping generation one.

[3] See Atkinson and Sandmo [2], Erosa and Gervais [11] and Garriga [12]; for a review see Renström [19] and Erosa and Gervais [10].

[4] In this model a crucial condition for the government to implement the "second best" policy is the availability of age-dependent taxes. The other central hypothesis, which is common to all the models mentioned above, is the presence of a "commitment technology", in order to guarantee the credibility of the capital taxation announced policy.

The aim of this work is to extend the analysis of optimal dynamic taxation by considering an OLG-perpetual youth (PY) model *à la* Blanchard [4] with growing population[5]. This extension enables us to encompass the issues mentioned above which, up to now, have been studied separately or under special assumptions. In fact by adopting the PY framework we can deal with overlapping generations, finite (expected) life-time horizon (via a constant instantaneous probability of death), life-cycle behavior and investigate the role played by both the intertemporal and intergenerational discount rates of the policymaker.

As known, the PY framework implies the existence of life insurance contracts, which individuals subscribe in order to offset the risk of dying; in this paper we allow also for a special kind of imperfection in the credit market, namely unfair life insurance contracts: this means that, in principle, individuals in each period can receive insurance payments that are different (and typically lower) than the actuarially fair ones.

The main results can be summarized as follows: first, similarly to the ILRA models, if the intertemporal elasticity of consumption is constant and the OLG mechanism is absent, the zero tax rule obtains even in the presence of a probability of dying.

Second, when the government is more (less) impatient than individuals (so that public and private intertemporal discount rates differ), the former finds it optimal to levy positive (negative) taxes due to Pigouvian correction motives: in fact, by doing this it lowers (increases) the current rhythm of over (under) accumulation of capital. However, when contrasting the benefits of such policies with the associated deadweight losses of distortionary taxation, we end up with an asymmetric result as for the long run: in fact, when the government is more patient, the current value of the distortion generated by capital income taxation tends to zero, so that it is still optimal to subsidize future consumption; on the contrary, when the government is less patient, such a distortion explodes to infinity, so that taxation must be zero. The asymmetry in the result is however ruled out in the special case of a logarithmic utility function, that displays a unitary intertemporal elasticity of substitution, since the change of future interest rates induced by the tax does not affect the planned consumption pattern[6].

---

[5]Buiter [5] and Weil [20] amend the Blanchard's model by allowing for population dynamics.

[6]The result is obtained by arguments equivalent to those in Lansing, even if the con-

Third, differently from the ILRA and similarly to the OLG-LC models, another source of non zero taxation in the long run stems from the dynamics of the intertemporal elasticity of consumption, which, in the PY framework is not necessarily constant at the steady state. However, even if this condition does not apply (and thus even in the absence of life cycle behavior) and the Pigouvian motive is ruled out, we find that taxation can be non zero if the government intergenerational discount factor changes through time. This element is obviously absent in ILRA models, but is also usually excluded in the OLG-LC framework, since up to now it has been taken as constant; such assumption can be well reasonable with finite lifetime horizon and invariant population (see, for instance, Erosa and Gervais [11]), since the share of each cohort is typically constant through time. But this is not the case in the PY framework if the birth rate is non zero, in that the demographic weights of each cohort decrease over time even if the net population growth rate is zero. Thus, while in the existing OLG-LC models the violation of the Judd result depends crucially on the life-cycle behavior of consumption, in the present work the mere existence of the OLG mechanism is sufficient for delivering it. In fact, we show that when agents are perpetually young, in the sense that the intertemporal allocation of consumption and leisure at each date does not depend on age, contrary to Erosa and Gervais [11] and Garriga [12], in this (special) case in which the life-cycle motive is ruled out, the non zero result applies. Moreover, such result holds even if age-dependent taxes are not available.

It is worth noting that the non zero result presented here sheds light on the scope of capital income taxation, which is the correction of suboptimality of the market allocation of an OLG economy, due to the disconnection between generations[7]. In fact, in absence of altruism, in each instant new individuals (i.e. the new born) get into the economy, whose welfare is not cared for by the existing generations.

Finally, we also show that the case of unfair life insurance contracts changes only the capital income tax level but not the qualitative result presented above.

The work proceeds as follows: in the first section we present the model and derive the equilibrium conditions for the decentralized economy. Next,

---

clusion is different, given that Lansing [17] assumes the equality between the public and private discount rates.

[7]The relevance of the disconnection has been firstly analyzed by Weil [20].

we characterize the Ramsey problem by adopting the primal approach. Finally, we present the results by focusing on the new ones. Concluding remarks and a technical appendix will end the work.

2 The model

We consider a neoclassical-production-closed economy in which there is a large number of agents and firms.

Private agents, who are identical in their preferences, differ as for their date of birth s; moreover they undergo a probability of dying in each period, equal to δ; since in each period there is also a fraction α of new born, the population growth rate is equal to $\alpha - \delta \equiv n$. As a consequence, a cohort of individuals born at date s, at time t has cardinality:

$$\alpha e^{-\delta t} e^{\alpha s} N(0)$$

with $N(0)$ the size of population at time 0 and $s \leq t$. Now, by setting $N(0)$ equal to one, without loss of generality, the size of the whole population, at time t, is:

$$N(t) = \int_{-\infty}^{t} \alpha e^{\alpha s - \delta t} ds = e^{nt}.$$

Furthermore, individuals offer labor and capital services to firms by taking the net-of-tax factor prices, $\widetilde{w}(s,t)$ and $\widetilde{r}(s,t)$ as given. Firms, which are identical to each other, own a constant return to scale technology F satisfying the Inada conditions and which transforms the factors into production-consumption units. Finally, the government can finance an exogenous and constant stream of public expenditure G, by issuing internal debt $B(t)$ and by raising proportional taxes both on interests and wages, referred to as $\tau^k(s,t)$ and $\tau^l(s,t)$ respectively. Notice that taxes can in principle be conditioned on the date of birth[8].

[8]This strong assumption can be ruled out if one eliminates life cycle behavior, since the individual growth rate of consumption is then independent of individual characteristics. Our results, in fact, do not rely on it.

2.1 Private agents

The agents' preferences can be represented by the following instantaneous utility function:

$$U\left(c\left(s,t\right),l\left(s,t\right)\right)$$

where $c\left(s,t\right)$ and $l\left(s,t\right)$ are instantaneous consumption and labor supply respectively of individuals of cohort s, as of instant t. Such utility function is strictly increasing in consumption and decreasing in labor, strictly concave, and satisfies the standard Inada conditions.

Agents maximize the (expected) discounted sum of lifetime utils by choosing the optimal time path of consumption (savings) and labor hours under the budget constraint.

That is:

$$\max_{\{c(t),l(t)\}_s^\infty} \int_s^\infty e^{-(\beta+\delta)(t-s)} U\left(c\left(s,t\right),l\left(s,t\right)\right) dt \tag{1}$$

$$sub\ \dot{a}\left(s,t\right) = \left(\widetilde{r}\left(s,t\right) + \widetilde{\delta}_r\left(s,t\right)\right) a\left(s,t\right) + \widetilde{w}\left(s,t\right) l\left(s,t\right) - c\left(s,t\right) \tag{2}$$

$$\lim_{t\to\infty} a\left(s,t\right) e^{-\int_s^t \left(\widetilde{r}(s,v)+\widetilde{\delta}_r(s,v)\right)dv} = 0, \quad a\left(s,s\right) = \overline{a}$$

where β is the intertemporal discount rate, a the agent's wealth; the notation $\dot{()}$ indicates the derivative with respect to time, while $\widetilde{r}\left(s,t\right) = r\left(t\right)\left(1-\tau^k\left(s,t\right)\right)$ and $\widetilde{w}\left(s,t\right) = w\left(t\right)\left(1-\tau^l\left(s,t\right)\right)$ are the net-of-tax factor prices. Notice that $\widetilde{\delta}_r$ is the instantaneous flow of income due to insurance (net of capital taxes)[9]; moreover, δ_r, the gross value, may differ from the actuarially fair value δ, due to market imperfections.

The FOCs of this problem imply:

$$U_{c(s,t)} = p\left(s,t\right) \tag{3}$$

$$U_{l(s,t)} = -p\left(s,t\right)\widetilde{w}\left(s,t\right) \tag{4}$$

$$-\left[\tilde{r}\left(s,t\right) + \widetilde{\delta}_r\left(s,t\right)\right] p\left(s,t\right) = \dot{p}\left(s,t\right) - \left(\beta+\delta\right) p\left(s,t\right) \tag{5}$$

[9] We assume here that the government taxes also life insurance payments; however, our results do not change qualitatively if this assumption is abandoned.

where the expression $U_{i(t)}$ is the partial derivative of the utility function with respect to argument $i = c, l$ at time t and $p(s,t)$ is the current value shadow price of wealth. According to such conditions, it can be shown that the growth rates of consumption and labor are:

$$\frac{\dot{c}}{c} = \left(\widetilde{r}(s,t) + \widetilde{\delta}_r(s,t) - (\beta + \delta)\right) \frac{1}{\theta_c} - \frac{\theta_{cl}}{\theta_c} \frac{\dot{l}}{l} \tag{6}$$

$$\frac{\dot{l}}{l} = \frac{1}{\theta_l} \frac{\left[\left(\widetilde{r}(s,t) + \widetilde{\delta}_r(s,t) - (\beta + \delta)\right)\left(1 - \frac{\theta_{lc}}{\theta_c}\right) - \frac{\dot{\widetilde{w}}(s,t)}{\widetilde{w}(s,t)}\right]}{1 - \frac{\theta_{cl}\theta_{lc}}{\theta_c \theta_l}}, \tag{7}$$

with $\theta_j = -\frac{U_{jj}j}{U_j}$, $j = c, l$, the elasticity of the marginal utility and $\theta_{ij} = -\frac{U_{ij}j}{U_i}$. Notice that, in case the utility function is additively separable in consumption and labor, the growth rates above are: $\frac{\dot{c}}{c} = \left(\widetilde{r}(s,t) + \widetilde{\delta}_r(s,t) - (\beta + \delta)\right) \frac{1}{\theta_c}$ and $\frac{\dot{l}}{l} = \left(\widetilde{r}(s,t) + \widetilde{\delta}_r(s,t) - (\beta + \delta)\right) \frac{1}{\theta_l}$.

2.2 Firms

Since firms run their business in a perfectly competitive framework, in each instant they hire capital and labor services according to their market prices (gross of taxes) and in order to maximize current period profits. This means that, for each firm i:

$$\frac{dF\left(K^i(t), L^i(t)\right)}{dK^i(t)} = r(t) \tag{8}$$

$$\frac{dF\left(K^i(t), L^i(t)\right)}{dL^i(t)} = w(t). \tag{9}$$

Due to the assumed identity of the firms and the presence of a CRS technology, such conditions can also be expressed for the economy as a whole, in per capita terms:

$$f_{k(t)} = r(t) \tag{8'}$$

$$f_{l(t)} = w(t), \tag{9'}$$

where $l(t) = \frac{L(t)}{N(t)} = \int_{-\infty}^{t} \nu_p(s,t)\, l(s,t)\, ds$, in which $\nu_p(s,t) = \alpha e^{-\alpha(t-s)}$ is the weight of cohort s in the whole population at period t.

2.3 The government and market clearing conditions

The government fixes an amount of exogenous public expenditure and finances it through taxes on income and by issuing debt. There is no constraint on the amount of debt (neither on the levels nor on the growth rates)[10]. We assume that the government has access to a commitment technology that prevents it from revising the announced path of distortionary tax rates whenever the possibility of lump sum taxation arises[11]. Thus, one obtains the usual condition:

$$\dot{B}(t) = r(t) B(t) + G - T(t). \tag{10}$$

Finally, since the market clearing condition implies that, at each date, the sum of capital and debt equal the aggregate private wealth, that is:

$$A(t) = K(t) + B(t), \tag{11}$$

then, eq. (10) can be also written as

$$\int_{-\infty}^{t} \alpha e^{\alpha s-\delta t} \Big[\dot{b}(s,t) - \Big(\widetilde{r}(s,t) + \widetilde{\delta}_r(s,t)\Big) b(s,t) + \tau^l(s,t) w(t) l(s,t)$$

$$+ (\delta_r - \delta) b(s,t) + \tau^k(s,t)(r(t) + \delta_r(s,t)) k(s,t) - g\Big] ds = 0. \tag{12}$$

3 The Ramsey problem

Since the primal approach to the Ramsey [18] problem consists in the maximization of a direct utility function through the choice of quantities (i.e. allocations)[12], a key point is restricting the set of allocations among which the

[10] The only constraint on the debt law of motion is the usual no-Ponzi game condition, namely: $\lim_{t\to\infty} B(t) e^{-\int_0^t r(v)dv} = 0$, and the starting condition $B(0) = \overline{B}$.

[11] This point concerns the "time inconsistency" problem affecting optimal taxation when a dynamic set up is considered: typically, the government has incentives to deviate from the announced (ex-ante) second best policy, upon achieving the instant in which the policy is phased in; in fact this happens because the stock of accumulated capital ex-post is perfectly rigid and now should be taxed more heavily, since this would mimic a lump sum taxation. The commitment hypothesis implies also that the capital tax at the beginning of the policy is given, that is, fixed exogenously at a level belonging to the $(0,1)$ interval.

[12] See Atkinson and Stiglitz [3]; on the other hand, the "dual" approach takes prices and tax rates as control variables (see Chamley [6] and Renström [19] for some examples).

government can choose to those that can be decentralized as a competitive equilibrium. Thus, in this paragraph we define a competitive equilibrium and the constraints that must be imposed to the policymaker problem, in order to achieve such a competitive outcome.

The first constraint can be obtained as follows: first, by taking eq. (2) and multiplying both sides by $e^{-\int_s^t \left[\widetilde{r}(s,v)+\widetilde{\delta}_r(s,v)\right]dv}$, we can write the following expression:

$$\frac{d\left[a(s,t)\, e^{-\int_s^t \left[\widetilde{r}(s,v)+\widetilde{\delta}_r(s,v)\right]dv}\right]}{dt} = e^{-\int_s^t \left[\widetilde{r}(s,v)+\widetilde{\delta}_r(s,v)\right]dv} \left[\widetilde{w}(s,t)\, l(s,t) - c(s,t)\right];$$

next, by multiplying both sides by $p(s,t)$ and exploiting the individuals' FOCs (3 to 5) we obtain:

$$p(s,s)\, e^{-\int_s^t \left[\widetilde{r}(s,v)+\widetilde{\delta}_r(s,v)-(\beta+\delta)\right]dv} \frac{d\left[a(s,t)\, e^{-\int_s^t \left[\widetilde{r}(s,v)+\widetilde{\delta}_r(s,v)\right]dv}\right]}{dt} =$$

$$-e^{-\int_s^t \left[\widetilde{r}(s,v)+\widetilde{\delta}_r(s,v)\right]dv} \left[U_l(s,t)\, l(s,t) + U_c(s,t)\, c(s,t)\right] \quad \Rightarrow$$

$$-U_c(s,s) \frac{d\left[a(s,t)\, e^{-\int_s^t \left[\widetilde{r}(s,v)+\widetilde{\delta}_r(s,v)\right]dv}\right]}{dt} = e^{-(\beta+\delta)(t-s)} \left[U_l(s,t)\, l(s,t) + U_c(s,t)\, c(s,t)\right];$$

finally, by integrating out and exploiting the individual's transversality condition, we get:

$$\int_s^\infty e^{-(\beta+\delta)(t-s)} \left[U_{c(s,t)} c(s,t) + U_{l(s,t)} l(s,t)\right] dt = a(s,s)\, U_{c(s,s)}, \ \forall s, \quad (13)$$

which is referred to as the "implementability constraint"[13].

As for the second constraint, writing eq. (2) in the following way:

$$\dot{a}(s,t) = \left[r(t) + \delta_r\right] a(s,t) + w(t)\, l(s,t) - c(s,t)$$

$$-\tau^k(s,t)\left[r(s,t) + \delta_r\right] a(s,t) - \tau^l(s,t)\, w(t) l(s,t); \quad (14)$$

integrating over the population to get the aggregate wealth:

[13]Such constraint must be satisfied $\forall s$. In the rest of the paper we assume for simplicity that $a(s,s) = \overline{a}$ is equal to zero for each cohort.

$$A(t) = \int_{-\infty}^{t} a(s,t)\,\alpha e^{-\delta t} e^{\alpha s} ds;$$

then, deriving with respect to time, one gets:

$$\dot{A}(t) = \underbrace{a(t,t)\,\alpha e^{-\delta t} e^{\alpha t}}_{=0} + \int_{-\infty}^{t} \frac{d\left[a(s,t)\,\alpha e^{-\delta t} e^{\alpha s}\right]}{dt} ds$$

where $a(t,t)$ is the initial wealth of individuals, which is supposed to be zero.

The expression above can be written as:

$$\dot{A}(t) = -\delta A(t) + \int_{-\infty}^{t} \dot{a}(s,t)\,\alpha e^{-\delta t} e^{\alpha s} ds, \tag{15}$$

so that, including (14) into (15), we obtain:

$$\dot{A}(t) = -\delta A(t) + [r(t) + \delta_r]\,A(t) - [r(t) + \delta_r] \int_{-\infty}^{t} \tau^k(s,t)\,a(s,t)\,\alpha e^{-\delta t} e^{\alpha s} ds + \tag{16}$$
$$- C(t) + W(t) - \int_{-\infty}^{t} \tau^l(s,t)\,w(t) l(s,t)\,\alpha e^{-\delta t} e^{\alpha s} ds,$$

where $C(t)$ and $W(t)$ are aggregate consumption and gross aggregate wages, respectively. Note that the sum of the two integrals in eq. (16) is the total amount of revenues, $T(t)$.

Finally, recalling the law of motion of aggregate debt, exploiting the market clearing condition and substituting the expression for $T(t)$ of (10) into (16), we get:

$$\dot{K}(t) = (\delta_r - \delta)(K(t) + B(t)) + r(t)\,K(t) + W(t) - C(t) - G, \tag{17}$$

which can also be written as:

$$\textstyle\int_{-\infty}^{t} \alpha e^{\alpha s - \delta t} \Big[\dot{k}(s,t) - (\delta + r(t))\,k(s,t) - w(t)\,l(s,t)$$

$$- (\delta_r - \delta) (b(s,t) + k(s,t)) + c(s,t) + g] \, ds = 0. \qquad (18)$$

Such expression is usually referred to as the "feasibility constraint".

We can now give the following definition:

Definition 1 *A competitive equilibrium is: a) an infinite sequence of policies* $\pi = \left\{\tau^k(s,t), \tau^l(s,t), b(s,t)\right\}_0^\infty$, *b) allocations* $\{c(s,t), l(s,t), k(s,t)\}_0^\infty$ *and c) prices* $\{w(t), r(t)\}_0^\infty$ *such that, at each instant t: b) satisfies eq.* (1) *subject to* (2)*, given a) and c); c) satisfies eq.* (8′) *and eq.* (9′)*; eqs.* (18) *and* (12) *are satisfied.*

Such allocations are often referred to as "implementable".

In the light of the definition given above, the following proposition holds:

Proposition 1 *An allocation is a competitive equilibrium if and only if it satisfies implementability and feasibility.*

Proof. The first part of the proposition is true by construction. The reverse (any allocation satisfying implementability and feasibility is a competitive equilibrium) is provided in Appendix A. ■

3.1 Solution

Let us suppose that the policy is introduced at the end of period t_0.

The problem the policymaker faces can be stated as follows[14]:

$$\max_{\{c(s,t),l(s,t),k(s,t)\}_0^\infty} \int_{-\infty}^{t} \mu_g \int_{\max(s,t_0)}^{\infty} e^{-\gamma_g(t-\max(s,t_0))} \{U(c(s,t), l(s,t)) +$$

$$\overline{\lambda} \left[U_{cc} c(s,t) + U_l l(s,t)\right]\Big\} \, dtds$$

$$sub \int_{-\infty}^{t} \mu_p(s,t) \Big[\dot{k}(s,t) - (\delta + r(t)) k(s,t) - w(t) l(s,t) +$$

$$- (\delta_r - \delta) (b(s,t) + k(s,t)) + c(s,t) + g] \, ds = 0, \ \forall t > t_0,$$

[14] We incorporate the implementability constraint into the maximand, λ being the multiplier on the implementability constraint.

$$\lim_{t\to\infty} k(s,t)\, e^{-\int_{\max(s,t_0)}^{t}\left(\widetilde{r}(s,v)+\widetilde{\delta}_r(s,v)\right)dv} = 0, \quad a(s,t_0) \quad given, \ \forall s$$

where $\mu_g(s,t)$ and γ_g are the weight that the government attaches to the generation born in year s and the government discount rate, respectively[15], $\mu_p = \alpha e^{\alpha s-\delta t}$ the size of cohort s and $\overline{\lambda}$ is the current value multiplier associated to the implementability constraint, defined as $\overline{\lambda}(t) = \lambda e^{(\gamma-(\beta+\delta))(t-\max(s,t_0))}$.

Now, by differentiating the feasibility constraint we get[16]:

$$c(s,t) = -\dot{k}(s,t) + (\delta + r(t))\, k(s,t) + w(t)\, l(s,t) + (\delta_r - \delta)(b(s,t) + k(s,t)) - g.$$

By substituting it into the problem and by applying the calculus of variations method, the problem can be stated as follows[17]:

$$\max_{\{l,k\}_0^\infty} \int_{-\infty}^{t} \mu_g \int_{\max(s,t_0)}^{\infty} e^{-\gamma_g(t-\max(s,t_0))} \left\{U\left(c\left(k,\dot{k}\right),l\right) + \overline{\lambda}\left[U_c c\left(k,\dot{k}\right) + U_l l\right]\right\} dtds.$$

Thus, the solution for k is[18]:

$$e^{-\gamma_g(t-\max(s,t_0))}\left\{\mu_g U_c\left[1+\overline{\lambda}(1+H_c)\right](r+\delta_r) - \gamma_g \mu_g U_c\left[1+\overline{\lambda}(1+H_c)\right] + \dot{\mu}_g U_c\left[1+\overline{\lambda}(1+H_c)\right]\right.$$

$$\left. +\mu_g\left(U_{cc}\dot{c} + U_{cl}\dot{l}\right)\left[1+\overline{\lambda}(1+H_c)\right] + \mu_g U_c \dot{\overline{\lambda}}(1+H_c) + \mu_g U_c \overline{\lambda}\dot{H}_c\right\} = 0 \tag{19}$$

---

[15]Note that, in principle, the former parameter may depend also on t. Moreover, we omit the government budget constraint since, by Walras' law, it is satisfied if the implementability and feasibility constraints hold.

[16]This step hinges on the assumption that $\dot{b}(s,t) - \left(\widetilde{r}(s,t) + \widetilde{\delta}_r(s,t)\right) b(s,t) + \tau^l(s,t)\, w(t)\, l(s,t) + (\delta_r - \delta)\, b(s,t) + \tau^k(s,t)(r(t) + \delta_r(s,t))\, k(s,t) - g = 0$.

This means that the public balance is divided into "generational" accounts, the dynamics of which is controlled by the government via labor and capital income taxes specific to each generation.

[17]From now onward, we omit both the s and t indexes, when it does not generate ambiguity.

[18]See Appendix B for the solution conditions of this problem. Note that the interiority of the solution is guaranteed by the Inada conditions. However, the FOCs are necessary but not sufficient due to the possible non convexity of the implementability constraint. The solution for l is omitted for brevity.

where the term $H_i = \frac{U_{ii}i + U_{ji}j}{U_i}$ is what is usually referred to as the "general equilibrium elasticity". Now, by dividing expression (19) by $U_c\mu_g\left[1+\overline{\lambda}\left(1+H_c\right)\right]$, and rearranging terms, we get:

$$\frac{\dot{c}}{c} = \frac{1}{\theta_c}\left\{\left(r+\delta_r-\gamma_g\right)+\frac{\dot{\mu}_g}{\mu_g}-\frac{\dot{\overline{\lambda}}\left(1+H_c\right)}{\left[1+\overline{\lambda}\left(1+H_c\right)\right]}+\frac{\overline{\lambda}\dot{H}_c}{\left[1+\overline{\lambda}\left(1+H_c\right)\right]}-\theta_{cl}\frac{\dot{l}}{l}\right\}.$$

Substituting for the growth rate of consumption stemming from the individual optimization condition (eq. (6)), we get the expression for the optimal capital income tax:

$$\tau^k = \frac{1}{f_k+\delta_r}\left\{\left[\gamma_g-(\beta+\delta)\right]-\frac{\dot{\overline{\lambda}}\left(1+H_c\right)}{\left[1+\overline{\lambda}\left(1+H_c\right)\right]}-\frac{\dot{\mu}_g}{\mu_g}-\frac{\overline{\lambda}\dot{H}_c}{\left[1+\overline{\lambda}\left(1+H_c\right)\right]}\right\}. \tag{20}$$

4 Discussion of the results

We now discuss the results concerning capital income taxation, in both the short and the long run.

Preliminarily, it is worth noting that eq. (20) does not yield an explicit formula for τ^k, since H_c depends upon the tax rate itself[19].

Next, eq. (20) shows that the imperfection in the insurance market does not determine whether the tax rate is different from zero or not, since it appears only in the denominator.

We can now state the following proposition:

Proposition 2 *If the economy converges to a steady state, along the transition path, for $t > 0$, the tax on capital income is in general different from zero unless $\dot{H}_c = 0$, $\dot{\mu}_g = 0$ and $\gamma_g = (\beta+\delta)$. At the steady state the capital income tax is different from zero, unless a) $\dot{H}_c = 0$, $\dot{\mu}_g = 0$ and $\gamma_g \geq \beta+\delta$ or b) $\gamma_g < \beta+\delta$ and $\left[\gamma_g-(\beta+\delta)\right] = \frac{\dot{\mu}_g}{\mu_g}$.*

[19]Moreover, we do not have any condition ensuring that the tax rate will be in the $(0,1)$ interval, while we would suspect capital taxes to get sticking at the interval boundary for a (finite) period of time since the introduction of the policy. However, in the rest of the work we maintain the assumption of interiority of the equilibrium tax rates, for $t > 0$.

Proof. As for the transition phase, the proof is straightforward by inspection of eq. (20), which, by substituting for $\dot{\overline{\lambda}}$, becomes:

$$\tau^k = \frac{1}{f_k + \delta_r}\left\{\frac{\left[\gamma_g - (\beta+\delta)\right]}{1+\overline{\lambda}\,(1+H_c)} - \frac{\dot{\mu}_g}{\mu_g} - \frac{\overline{\lambda}\dot{H}_c}{\left[1+\overline{\lambda}\,(1+H_c)\right]}\right\}.$$

As for the steady state, to better understand the implications of the model, we distinguish three cases, according to whether the policymaker discount rate γ_g is equal, higher or lower than the individual one.

1. $\gamma_g = \beta+\delta$. In this case $\overline{\lambda} \to \lambda$, so that $\tau^k = \frac{1}{f_k+\delta_r}\left\{-\frac{\dot{\mu}_g}{\mu_g} - \frac{\lambda \dot{H}_c}{[1+\lambda(1+H_c)]}\right\}$.
2. $\gamma_g > \beta+\delta$. In this case $\overline{\lambda} \to \infty$, and $\tau^k = \frac{1}{f_k+\delta_r}\left\{-\frac{\dot{\mu}_g}{\mu_g} - \frac{\dot{H}_c}{(1+H_c)}\right\}$.
3. $\gamma_g < \beta+\delta$. $\overline{\lambda} \to 0$ and $\tau^k = \frac{1}{f_k+\delta_r}\left[\left[\gamma_g - (\beta+\delta)\right] - \frac{\dot{\mu}_g}{\mu_g}\right]$.

■

From the proposition above, it emerges that there are four independent forces determining the level of τ^k: 1) the dynamics of H_c $\left(\dot{H}_c\right)$; 2) the difference between the government (γ_g) and individual ($\beta+\delta$) intertemporal discount factors; this also determines the third factor, i.e. 3) the dynamics of $\overline{\lambda}$, the multiplier on the implementability constraint; 4) the dynamics of the social intergenerational weight ($\dot{\mu}_g$); .We can now briefly comment the role of the factors determining the optimal tax rate.

4.1 The role of the utility function

Factor 1) has been widely discussed in the literature: $\dot{H}_c = 0$ obtains if one assumes that the utility function is homothetic in consumption and (weakly) separable in consumption and leisure. Otherwise, future consumption is taxed/subsidized if consumption demand is getting more/less inelastic. Moreover, as recalled above, this factor marks the difference between the ILRA and the OLG-LC models as for the steady state result: in fact, in OLG models (and also in the PY version) H_c can vary with age even at the steady state. However, as shown in eq. (20), even in the absence of a life cycle, in the present model the non zero tax rule can still apply.

4.2 The role of the time discounting rates

Factors 2) and 3) can be discussed together, since they are both dependent on the relationship between the government and individual time discount rates. For simplicity, let us assume that $\dot{\mu}_g$and $\dot{H}_c$are equal to zero. If $\gamma_g = \beta + \delta$ (or $\delta = 0$ and $\gamma_g = \beta$), then there is no scope for capital income taxation, either along the transition path or at the steady state (see Erosa and Gervais [11]). The difference between γ_g and $(\beta + \delta)$, instead, opens the way to a Pigouvian correction. For the case of $\beta + \delta > \gamma$, individuals are discounting the future at a rate that is higher than the government one. As a consequence, since they are consuming at a too high rate, the government finds it optimal to subsidize capital, that is, future consumption. The same reasoning, with opposite conclusions, applies in the second case $(\beta + \delta < \gamma)$. In both situations, however, it is worth noting that the tax rate is inversely proportional $\overline{\lambda}$, which is usually interpreted as the (current value of the distortion) brought about by non lump-sum taxation. As expected, this relationship shown that, in general, increasing the capital income tax worsens the overall deadweight loss. However, in the first case this parameter tends to zero, while it gets bigger towards infinity in the other case. As a consequence, in the long run, when $\overline{\lambda}$ decreases through time, there is still room for subsidizing future consumption (i.e. current capital income). On the other hand, when $\overline{\lambda}$ raises exponentially to infinity, the distortionary effect overwhelms the welfare improvement due to the Pigouvian correction. Hence, the government, which cares relatively less about the present than individuals do, finds it optimal to announce a zero capital income tax for the long run. The only exception for such asymmetry in the result is the case of a logarithmic utility function, with $H_c = -1$. In fact, since the utility function displays a unitary intertemporal elasticity of substitution, the substitution and the income effect generated by an interest rate variation (due to taxation) cancel out. Hence, the change of future interest rates, that is, the change of the relative prices of future consumption, does not distort the individual consumption/saving allocation. As a consequence, the long run tax can be positive and equal to $\frac{\gamma_g - (\beta + \delta)}{f_k + \delta_r}$.

4.3 The role of the intergenerational discount factor

As for factor 4), its role can be isolated by supposing $\gamma = (\beta + \delta)$ and $\dot{H}_c = 0$. Then, $\tau^k = -\frac{\frac{\dot{\mu}_g}{\mu_g}}{f_k + \delta_r}$, which is different from zero provided that μ_g, the weight assigned to each cohort by the government, is constant through time. This assumption is made in the existing OLG-LC models (see, in particular, Erosa and Gervais [11] and Garriga [12]: this element, together with the equality between the government and the individual intertemporal discount rates, leads to zero taxation in the absence of life cycle, i.e. $\dot{H}_c = 0$. However, this is typically not the case in this set up. Let us consider the situation in which the social intergenerational weight is equal to the actual demographic weight of each cohort within the population, i.e. $\mu_g = \frac{\mu_p}{e^{nt}}$. Given our assumption of a constant probability of death δ and a constant birth rate α, the relative size of each cohort is decreasing through time, so that $\frac{\dot{\mu}_g}{\mu_g} = \alpha$ for each cohort and, hence, the tax is $\tau^k = \frac{\alpha}{f_k + \delta_r}$. This new result clarifies that an independent source of taxation is represented by the "disconnection" of the economy, which is typical of OLG models: in fact, given the dynamics of μ_g, the government discriminates future consumption in favour of the present one; under a different perspective, one can note that at each date individuals tend to oversave relatively to what could maximize welfare, since they do not take into consideration the new born of the economy. In fact these individuals at each date provide extra resources for redistributing (at least the burden of taxation) among the existing generations (it is easy to show that, if individuals were altruistic towards their descendants, the JC zero tax rule would be restored).

Moreover, at least at the steady state, such tax does not depend on age. Erosa and Gervais have shown that, when age-dependent taxes are not available, the zero tax result does not generally apply in OLG models, since new constraints featuring the problem are violated (that is, the equality of the marginal rates of substitution of consumption among individuals of different cohorts). In our case, however, it is easy to show that such constraints (which take the form of the equality of consumption growth rates among individuals) are satisfied in the absence of life-cycle behavior (i.e. when $\dot{H}_c = 0$)[20].

As for the case of $\gamma_g \neq \beta + \delta$,when $\gamma_g > \beta + \delta$, along the transition path both factors drive to a positive taxation of capital income. As for individual consumption, the effects of such a policy are, *ceteris paribus*, to lower its growth rate with respect to that obtaining without taxation. On the other

[20]This property of the model can be verified by looking at eq. 6, which is constant for each individual if θ_c and θ_{cl} are constant among individuals, and by reckoning that $H_c = -(\theta_c + \theta_{cl})$.

hand, in case $\gamma_g < \beta+\delta$, there is a contrasting force at work: in fact, since the government is more forward looking (i.e. less impatient) than individuals, it tends to subsidize future consumption; therefore, the sign of the tax will depend on which force prevails. As for the steady state, in the first case factor 1) becomes irrelevant because the cost of using distortionary taxation tends to infinity. In the second one, instead, $\tau^k = \frac{1}{f_k+\delta_r}\left(\gamma_g - \beta - \delta + \alpha\right)$, which is zero only if $\gamma_g = \beta - n$. As a consequence, the zero tax result emerges in two very special cases, i.e. when $\gamma_g = (\beta - n)$ or $\gamma_g =(\beta + \delta)$ and $\alpha = 0$.

5 Conclusions

We tackle the issue of taxing capital income in a perpetual youth model *à la* Blanchard (i.e. an overlapping generation framework with individuals facing a constant probability of dying) by applying the primal approach to the Ramsey problem. Although less handleable than the traditional ones, this extension enables us to provide a more general model which, on the one hand, encompasses most of the existing results obtained in separated frameworks, and, on the other hand, delivers new insights.

The thrust of the paper is that several forces are at work leading to a non zero tax rate, in both the short and the long run.

Namely, we unveil the presence of four forces: a) the dynamics of the general equilibrium elasticity of consumption (H_c); b) the difference between the government and individual intertemporal discount rates, which also determines c) the dynamics of the distortionary cost of taxation as for the government; d) the dynamics of the social intergenerational discount rate; in the case this corresponds to the actual share of each cohort within the population, this leads to a positive taxation of capital income proportionally to the birth rate.

The first factor has been widely discussed in the literature: given the equivalence between capital income and future consumption taxation, it is convenient to hit the latter relatively more if the intertemporal elasticity of consumption is decreasing. The economic intuition underlying the role of factor b) and c) is the following: the different degree of patience between the policy maker and individuals generates an incentive for the former to levy positive or negative taxes on capital for a Pigouvian correction; precisely, when the policymaker is less (more) patient than individuals, it finds it optimal to levy a positive (negative) tax on capital income, so as to lower (increase) the current consumption growth rate. However, the benefits of such policy must be confronted with the cost of using distortionary taxation. In fact, the positive taxation policy leads to an explosive distortionary effect, which prevents the government to implement it. The only case in which such distortionary effect does not play any role is that of unitary elasticity of substitution of consumption (i.e. logarithmic utility): in fact, in this case future change of the interest rate caused by policy does not distort the consumption pattern chosen by individuals (since the income and substitution effects cancel out): hence, the government has room again for positive taxation of interest income.

Finally, factor d) while being absent in both ILRA models and in OLG ones with constant population, plays a crucial role when, as it is likely to happen in the real world, the dynamics of the population is more complicated. We show that, in a PY framework, both the size and the demographic weight of cohorts decrease, so that a varying social intergenerational discount rate appears a sensible rather than an ad hoc assumption. In fact, when the latter equals the actual share of each generation, a positive tax on capital income turns out to be optimal. This occurs because individuals, who are disconnected to each other, do not take into account the fact that the arrival of the new born, at each date provides extra resources for an intergenerational redistribution of the burden of taxation. On the other hand, by reckoning this possibility, the government reduces the oversaving of individuals by hitting future consumption proportionally to the birth rate of the economy.

Concluding, from the analysis above it turns out that the violation of the Chamley Judd rule does crucially depend upon the assumption of OLG (without altruism) and/or finite lifetime horizon. In fact such devices generate a difference between the optimal rate of individual consumption growth and that resulting in the absence of taxation, which thus gives room to corrective public intervention. However, differently from the existing OLG-LC models, the presence of life cycle behavior is not a necessary condition for the non zero tax result, which obtains even if the government and individual discount rates are equal and the general equilibrium elasticity of consumption is constant.

References

[1] A. Atkeson, V.V. Chari, P.J. Kehoe, Taxing Capital Income: A Bad Idea, Fed. Reserve Bank Minneapolis Quart. Rev. **23** (1999), 3-17.

[2] A.B Atkinson, A. Sandmo, Welfare Implications of the Taxation of Savings, Econ. J. **90** (1980), 529-549.

[3] A.B. Atkinson, J.E. Stiglitz, Lectures on public economics, McGraw-Hill, London, 1980.

[4] O.J. Blanchard, Debt, Deficits, and Finite Horizons, J. Polit. Economy **93** (1985), 223-247.

[5] W.H. Buiter, Death, Birth, Productivity Growth and Debt Neutrality, Econ. J. **98** (1988), 279-293.

[6] C. Chamley, Optimal taxation of capital income in general equilibrium with infinite lives, Econometrica, **54** (1986), 607-622.

[7] V.V. Chari, P.J. Kehoe, Optimal Fiscal and Monetary Policy, in. J.B. Taylor, M. Woodford (Eds.), Handbook of Macroeconomics, Vol. 1, North-Holland, Amsterdam, 1999, pp1670-1745.

[8] W.J. Corlett, D.C. Hague, Complementarity and the excess burden of taxation, Rev. Econ. Stud. **21** (1953), 21-30.

[9] V. De Bonis, L. Spataro, Taxing capital income: is it really a bad idea?, Studi e Ricerche, Dipartimento di Scienze Economiche, Università degli Studi di Pisa, n.88, 2002.

[10] A. Erosa, M. Gervais, Optimal Taxation in Infinitely-Lived Agent and Overlapping Generations Models: A Review, Fed. Reserve Bank Richmond Econ. Quart. **87** (2001), 23-44.

[11] A. Erosa, M. Gervais, Optimal Taxation in Life-Cycle Economies, J. Econ. Theory **105** (2002), 338-369.

[12] C. Garriga

[13] L.E. Jones, R.E. Manuelli, P.E. Rossi, On the Optimal Taxation of Capital Income, J. Econ. Theory **73** (1997), 93-117.

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

[15] K.L. Judd, Optimal Taxation and Spending in General Competitive Growth Models, J. Public Econ. **71** (1999), 1-26.

[16] M.I. Kamien, N.L. Schwartz, Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management, Second Edition, North Holland, New York, 1991.

[17] K.J. Lansing, Optimal redistributive capital taxation in a neoclassical growth model, J. Public. Econ. **73** (1990), 423-453.

[18] F.P. Ramsey, A contribution to the theory of taxation, Econ. J. **37** (1927), 47-61.

[19] T. Renström, Optimal Dynamic Taxation, in: S.B. Dahiya (Ed.), The Current State of Economic Science, Vol. 4, Spellbound Publications, 1999, pp1717-1743.

[20] P. Weil, Overlapping families of infinitely-lived agents, J. Public Econ. **38** (1989), 183-198.

[21] I. Yakadina, Optimal Capital-Labor Taxes under Uncertainty and Limits on Debt, Universitat Pompeu Fabra, mimeo, 2001.

[22] X. Zhu, Optimal Fiscal Policy in a Stochastic Growth Model, J. Econ. Theory **58** (1992), 250-289.

6 Appendix A: Proof of Proposition 1

Proof. Since a competitive equilibrium (or implementable allocation) satisfies both the feasibility and the implementability constraints by construction, in this Appendix we demonstrate the reverse of Proposition 1: any feasible allocation satisfying implementability is a competitive equilibrium.

Suppose that an allocation satisfies the implementability and the feasibility constraints. Then, define a sequence of after tax prices as follows: $\widetilde{w}(s,t) = -\frac{U_{l(s,t)}}{U_{c(s,t)}}$, $\left[\tilde{r}(s,t) + \widetilde{\delta}_r(s,t)\right] = \left(\beta + \delta - \frac{\dot{p}(s,t)}{p(s,t)}\right)$, with $p(s,t) = U_{c(s,t)}, \forall s$ and $\forall t$, and a sequence of before tax prices as: $f_{k(t)} = r(t)$ and $f_{l(t)} = w(t)$. As a consequence, by construction such allocation satisfies both the consumers' and firms' optimality conditions.

The second step is to show that the allocation satisfies the consumer budget constraint. Take the implementability constraint and substitute $U_{c(s,t)}$, $U_{l(s,t)}$ by using the expressions above:

$$\int_s^{\infty} e^{-(\beta+\delta)(t-s)} \left[p(s,t)\, c(s,t) - \widetilde{w}(s,t)\, p(s,t)\, l(s,t)\right] dt = a(s,s)\, p(s,s), \quad \forall s$$

then, by exploiting the expression for $\dot{p}(s,t)$ we get[21]:

$$\int_s^{\infty} p(s,s) e^{-(\beta+\delta)(t-s)} e^{-\int_s^t \left[\widetilde{r}(s,v)+\widetilde{\delta}_r(s,v)-(\beta+\delta)\right]dv} \left[c(s,t) - \widetilde{w}(s,t)\, l(s,t)\right] dt = a(s,s)\, p(s,s).$$

Finally, by eliminating $p(s,s)$ and defining $c(s,t) - \widetilde{w}(s,t)\, l(s,t) = \widetilde{r}(s,t)\, q(s,t) - \dot{q}(s,t)$ we get:

$$-\int_s^{\infty} \frac{d\left[q(s,t)\, e^{-\int_s^t \left[\widetilde{r}(s,v)+\widetilde{\delta}_r(s,v)\right]dv}\right]}{dt} dt = a(s,s)$$

which holds if $q(s,t) = a(s,t)$ and $\lim_{t\to\infty} a(s,t)\, e^{-\int_0^t \left[\widetilde{r}(s,v)+\widetilde{\delta}_r(s,v)\right]dv} = 0$.

Finally, as for the public sector budget constraint, by substituting the expression for consumption obtainable by the individual budget constraint into the feasibility constraint, we get:

$$\int_{-\infty}^{t} \alpha e^{\alpha s-\delta t} \Big[\dot{k}(s,t) - (\delta + r(t))\, k(s,t) - w(t)\, l(s,t) - (\delta_r - \delta)\left(b(s,t) + k(s,t)\right)$$
$$-\dot{a}(s,t) + \left(\widetilde{r}(s,t) + \widetilde{\delta}_r(s,t)\right) a(s,t) + \widetilde{w}(s,t)\, l(s,t) + g\Big]\, ds = 0.$$

Finally, by defining $b(t) = k(t) - a(t)$ and exploiting the definition of taxes, the previous expression becomes:

$$\int_{-\infty}^{t} \alpha e^{\alpha s-\delta t} \Big[\dot{b}(s,t) - \left(\widetilde{r}(s,t) + \widetilde{\delta}_r(s,t)\right) b(s,t) + \tau^l(s,t)\, w(t)\, l(s,t) - g$$
$$+ (\delta_r - \delta)\, b(s,t) + \tau^k(s,t)\left(r(t) + \delta_r(s,t)\right) k(s,t)\Big]\, ds = 0,$$

which is eq. (12) in the text. ■

[21] The equations below hold $\forall s$.

7 Appendix B: The "calculus of variations" method

We now sketch the strategy adopted for solving the Ramsey problem presented in Section 3.1.

Following Kamien and Schwartz [16], suppose the problem has the form

$$\max \int \int F\left(t, s, x\left(t, s\right), x_t\left(t, s\right), x_s\left(t, s\right)\right) dsdt$$

where the symbol x_y indicate the partial derivatives of variable x with respect to y (x can be also a vector of variables). The Euler equation for such a problem is the following:

$$F_x - \partial F_{x_t}/\partial t - \partial F_{x_s}/\partial s = 0.$$

Moreover, in case the problem contains also a (double) integral constraint, such as:

$$\int \int q\left(t, s, x\left(t, s\right), x_t\left(t, s\right), x_s\left(t, s\right)\right) dsdt = 0,$$

this constraint can be appended to the integrand with a multiplier function $\lambda\left(t, s\right)$, so that, if the solution x^* maximizing F subject to the constraint does exist, then there is a function $\lambda\left(t, s\right)$ such that x^* satisfies the Euler equations for

$$\int \int \left(F + \lambda q\right) dsdt.$$