

Taxing capital income in a perpetual youth economy

Valeria De Bonis¹

Dipartimento di Scienze Economiche, Università di Pisa, Italy
debonis@ec.unipi.it

Luca Spataro²

Dipartimento di Scienze Economiche, Università di Pisa, Italy
l.spataro@ec.unipi.it

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

²Tel.: +39 050 2216333, fax: +39 050 598040.

Abstract

We reconsider the issue of capital income taxation in a perpetual youth framework. We show that the zero tax result does not generally hold, even in the long-run and in the presence of homothetic in consumption and separable preferences. In fact, at the steady state, there are at least three forces pushing toward the taxation of capital income: the probability of death combined with the overlapping generation mechanism, the difference between the weight attached to each generation by the government and its demographic weight and the relationship between the government and the individual intertemporal discount rates. Finally, we show 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 [13] 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 [14] 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: 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 concerned¹, 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 authors³, in this setup a non zero tax rate result holds in general, even in the long run, since optimal consumption and labor are not constant over life, in the presence of life-cycle behavior⁴.

Finally, another source of non zero taxation, highlighted in both ILRA and OLG-LC models, stems from the difference between government and

¹See Atkeson et al. [1] and Chari et al. [7].

²Both 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.

³See Atkinson and Sandmo [2] and Erosa and Gervais [11]; for a review see Renström [17] and Erosa and Gervais [10].

⁴In 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 announced capital income taxation policy.

individual discount rates. De Bonis and Spataro [9], for example, end up with a non zero (negative) tax on capital income, even in the long run and with homothetic in consumption and separable utility functions, if the government is more patient than individuals, while the Chamley-Judd result is still valid if the government is less patient⁵.

The aim of this work is to extend the analysis of optimal dynamic taxation by considering a perpetual youth (PY) model *à la* Blanchard [4] with growing population⁶. 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 life-time horizon (via a constant probability of death), life-cycle behavior and investigate the role played by both the intertemporal and intergenerational discount rates of the policymaker. Another feature by which we depart from the previous literature is to allow for a special kind of imperfection in the credit market, namely unfair life insurance contracts.

The main results can be summarized as follows: first, the zero tax rule is violated, in both the short and the long run, in the presence of the (assumed) difference between the weight that the government attaches to each generation and each generation's actual weight in the current population; second, the same failure occurs in the presence of the probability of dying *and* the OLG framework, even if the government and individual's intertemporal/intergenerational discount rates are equal. Third, we show that the presence of unfair life insurance contracts influences only the level but not the qualitative result of the non zero taxation. In the light of these findings, the Chamley-Judd result turns out to be a special case.

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

⁵Among other articles focusing on the optimal capital income taxation problem see Jones et al. [12], modeling human capital accumulation, and Chari et al. [7], Zhu [20] and Yakadina [19], dealing with stochastic frameworks.

⁶Buiter [5] and Weil [18] amend the Blanchard's model by allowing for population dynamics.

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, $\tilde{w}(s, t)$ and $\tilde{r}(s, t)$ 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 birth⁷.

2.1 Private agents

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

$$U(c(s, t), l(s, t))$$

⁷This strong assumption can be ruled out if one eliminates life cycle behavior. Our results, in fact, do not rely on it.

where $c(s, t)$ and $l(s, t)$ 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:

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

$$sub \dot{a}(s, t) = (\tilde{r}(s, t) + \tilde{\delta}_r(s, t)) a(s, t) + \tilde{w}(s, t) l(s, t) - c(s, t) \quad (2)$$

$$\lim_{t \rightarrow \infty} a(s, t) e^{-\int_s^t (\tilde{r}(s, v) + \tilde{\delta}_r(s, v)) dv} = 0, \quad a(s, s) = \bar{a}$$

where β is the intertemporal discount rate, a the agent's wealth; the notation $\dot{(\)}$ indicates the derivative with respect to time, while $\tilde{r}(s, t) = r(t) (1 - \tau^k(s, t))$ and $\tilde{w}(s, t) = w(t) (1 - \tau^l(s, t))$ are the net-of-tax factor prices. Notice that $\tilde{\delta}_r$ is the instantaneous flow of income due to insurance (net of capital taxes)⁸; moreover, δ_r , the gross value, may differ from the actuarially fair value δ , due to market imperfections.

The FOCs of this problem imply:

$$U_{c(s, t)} = p(s, t) \quad (3)$$

$$U_{l(s, t)} = -p(s, t) \tilde{w}(s, t) \quad (4)$$

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

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:

⁸We assume here that the government taxes also life insurance payments; however, our results do not change qualitatively if this assumption is abandoned.

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

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

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

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 :

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

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

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:

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

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

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)⁹. 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 arises¹⁰. Thus, one obtains the usual condition:

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

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

$$A(t) = K(t) + B(t), \quad (11)$$

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

$$\int_{-\infty}^t \alpha e^{\alpha s - \delta t} \left[\dot{b}(s, t) - \left(\tilde{r}(s, t) + \tilde{\delta}_r(s, t) \right) b(s, t) + \tau^l(s, t) w(t) l(s, t) \right. \\ \left. + (\delta_r - \delta) b(s, t) + \tau^k(s, t) (r(t) + \delta_r(s, t)) k(s, t) - g \right] ds = 0. \quad (12)$$

3 The Ramsey problem

Since the primal approach to the Ramsey [16] problem consists in the maximization of a direct utility function through the choice of quantities (i.e. allocations)¹¹, a key point is restricting the set of allocations among which the government can choose to those that can be decentralized as a competitive equilibrium. Thus, in this paragraph we define a competitive equilibrium

⁹The only constraint on the debt law of motion is the usual no-Ponzi game condition, namely: $\lim_{t \rightarrow \infty} B(t) e^{-\int_0^t r(v) dv} = 0$, and the starting condition $B(0) = \bar{B}$.

¹⁰This 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.

¹¹See Atkinson and Stiglitz [3]; on the other hand, the “dual” approach takes prices and tax rates as control variables (see Chamley [6] and Renström [17] for some examples).

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 $e^{-\int_s^t [\tilde{r}(s,v) + \tilde{\delta}_r(s,v)] dv}$, we can write the following expression:

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

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

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

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

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

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

Since this constraint has to be satisfied for the whole economy, it must be

$$\int_{-\infty}^t \int_s^\infty \alpha e^{-\delta t} e^{\alpha s} \left\{ e^{-(\beta + \delta)(t-s)} [U_{c(s,t)} c(s, t) + U_{l(s,t)} l(s, t)] - e^{-(t-s)} a(s, s) U_{c(s,s)} \right\} dt ds = 0, \quad (14)$$

which is referred to as the “implementability constraint”¹².

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

¹²In the rest of the paper we assume for simplicity that $a(s, s) = \bar{a}$ is equal to zero for each cohort.

$$\begin{aligned} \dot{a}(s, t) &= [r(t) + \delta_r] a(s, t) + w(t) l(s, t) - c(s, t) \\ &\quad - \tau^k(s, t) [r(s, t) + \delta_r] a(s, t) - \tau^l(s, t) w(t) l(s, t); \end{aligned} \quad (15)$$

integrating over the population to get the aggregate wealth:

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

then, deriving with respect to time, one gets:

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

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

The expression above can be written as:

$$\dot{A}(t) = -\delta A(t) + \int_{-\infty}^t \dot{a}(s, t) \alpha e^{-\delta t} e^{\alpha s} ds, \quad (16)$$

so that, including (15) into (16), we obtain:

$$\begin{aligned} \dot{A}(t) &= -\delta A(t) + [r(t) + \delta_r] A(t) - [r(t) + \delta_r] \int_{-\infty}^t \tau^k(s, t) a(s, t) \alpha e^{-\delta t} e^{\alpha s} ds + \\ &\quad - C(t) + W(t) - \int_{-\infty}^t \tau^l(s, t) w(t) l(s, t) \alpha e^{-\delta t} e^{\alpha s} ds, \end{aligned} \quad (17)$$

where $C(t)$ and $W(t)$ are aggregate consumption and gross aggregate wages, respectively. Note that the sum of the two integrals in eq. (17) 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 (17), we get:

$$\dot{K}(t) = (\delta_r - \delta) (K(t) + B(t)) + r(t) K(t) + W(t) - C(t) - G, \quad (18)$$

which can also be written as:

$$\int_{-\infty}^t \alpha e^{\alpha s - \delta t} \left[\dot{k}(s, t) - (\delta + r(t)) k(s, t) - w(t) l(s, t) - (\delta_r - \delta) (b(s, t) + k(s, t)) + c(s, t) + g \right] ds = 0. \quad (19)$$

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 $\pi = \{\tau^k(s, t), \tau^l(s, t), b(s, t)\}_0^\infty$, 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. (19) 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 is the following:

$$\begin{aligned} & \max_{\{c(s,t), l(s,t), k(s,t)\}_0^\infty} \int_{\max(s, t_0)}^\infty \int_{-\infty}^t \mu_g(s, t) e^{-\gamma_g(t - \max(s, t_0))} U(c(s, t), l(s, t)) ds dt \\ & \text{sub } \int_{\max(s, t_0)}^\infty \int_{-\infty}^t \mu_p(s, t) \left\{ e^{-(\beta + \delta)(t - \max(s, t_0))} [U_{c(s, t)} c(s, t) + U_{l(s, t)} l(s, t)] + \right. \\ & \quad \left. - e^{-(t - \max(s, t_0))} a(s, \max(s, t_0)) U_{c(s, \max(s, t_0))} \right\} ds dt = 0 \end{aligned}$$

$$\text{and } \int_{-\infty}^t \mu_p(s, t) \left[\dot{k}(s, t) - (\delta + r(t))k(s, t) - w(t)l(s, t) + \right. \\ \left. - (\delta_r - \delta)(b(s, t) + k(s, t)) + c(s, t) + g \right] ds = 0, \quad \forall t > t_0,$$

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

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, respectively¹³, and $\mu_p = \alpha e^{\alpha s - \delta t}$ the size of cohort s .

Now, by differentiating the feasibility constraint we get:

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

By substituting it into the problem, we get¹⁴:

$$\max_{\{l, k\}_{\max(s, t_0)}^{\infty}} \int_{\max(s, t_0)}^{\infty} \int_{-\infty}^t \mu_g e^{-\gamma_g(t - \max(s, t_0))} U(c(k, \dot{k}), l) ds dt \\ \text{sub } \int_{\max(s, t_0)}^{\infty} \int_{-\infty}^t \mu_p \left\{ e^{-(\beta + \delta)(t - \max(s, t_0))} [U_c c + U_l l] + \right. \\ \left. - e^{-(t - \max(s, t_0))} a(s, \max(s, t_0)) U_{c(s, \max(s, t_0))} \right\} ds dt = 0.$$

By applying the calculus of variations method, the problem can be stated as follows:

$$\max_{\{l, k\}_{\max(s, t_0)}^{\infty}} \int_{\max(s, t_0)}^{\infty} \int_{-\infty}^t \left\{ \mu_g e^{-\gamma_g(t - \max(s, t_0))} U(c(k, \dot{k}), l) + \right. \\ \left. + \hat{\lambda} \mu_p \left[(U_c c + U_l l) - e^{(\beta + \delta - 1)(t - \max(s, t_0))} a(s, \max(s, t_0)) U_{c(s, \max(s, t_0))} \right] \right\} ds dt$$

¹³Note 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.

¹⁴From now onward, we omit both the s and t indexes, when it does not generate ambiguity.

where $\hat{\lambda}$ is the current value multiplier associated to the implementability constraint, defined as $\hat{\lambda}(t) = \lambda e^{-(\beta+\delta)(t-\max(s,t_0))}$. Thus, the solution for k is¹⁵:

$$e^{-\gamma_g(t-\max(s,t_0))} \left\{ U_c \mu_p \left[\frac{\mu_g}{\mu_p} + \bar{\lambda}(1+H_c) \right] \left[(r+\delta_r) - \gamma_g + (U_{cc}\dot{c} + U_{cl}\dot{l}) \right] + \right. \quad (20)$$

$$\left. U_c \left[\dot{\mu}_g + (1+H_c) \left(\bar{\lambda}\dot{\mu}_p + \dot{\bar{\lambda}} \right) + \mu_p \bar{\lambda} \dot{H}_c \right] \right\} = 0$$

where $\bar{\lambda} = \hat{\lambda} e^{\gamma_g(t-\max(s,t_0))} = \lambda e^{-(\beta+\delta-\gamma_g)(t-\max(s,t_0))}$ and 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 (20) by $U_c \mu_p$, and rearranging terms, we get:

$$\begin{aligned} \frac{\dot{c}}{c} = & \frac{1}{\theta_c} \left[(r+\delta_r - \gamma_g) - \delta \frac{\left[-\frac{\dot{\mu}_g}{\delta\mu_p} + \bar{\lambda}(1+H_c) \right]}{\left[\frac{\mu_g}{\mu_p} + \bar{\lambda}(1+H_c) \right]} + \right. \\ & \left. - (\beta+\delta-\gamma_g) \frac{\bar{\lambda}(1+H_c)}{\left[\frac{\mu_g}{\mu_p} + \bar{\lambda}(1+H_c) \right]} + \frac{\bar{\lambda}\dot{H}_c}{\left[\frac{\mu_g}{\mu_p} + \bar{\lambda}(1+H_c) \right]} - \theta_{cl} \frac{\dot{l}}{l} \right]. \end{aligned}$$

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:

$$\tau^k = \frac{1}{f_k + \delta_r} \left\{ (\gamma_g - (\beta+\delta)) + \frac{\delta \left[-\frac{\dot{\mu}_g}{\delta\mu_p} + \bar{\lambda}(1+H_c) \right] + (\beta+\delta-\gamma_g) \bar{\lambda}(1+H_c) - \bar{\lambda}\dot{H}_c}{\left[\frac{\mu_g}{\mu_p} + \bar{\lambda}(1+H_c) \right]} \right\}. \quad (21)$$

4 Discussion of the results

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

¹⁵See 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.

Preliminarily, it is worth noting that eq. (21) does not yield an explicit formula for τ^k , since H_c depends upon the tax rate itself¹⁶.

Next, eq. (21) 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.

Furthermore, there are four independent forces determining the level of τ^k : 1) the dynamics of H_c (\dot{H}_c); 2) the difference between the social intergenerational weight (μ_g) and the one corresponding to the size of each cohort (μ_p); 3) the difference between the government (γ_g) and individual ($\beta + \delta$) intertemporal discount factors; 4) the finite horizon deriving from the probability of death δ and the OLG mechanism.

Since factor 1) has been widely discussed in the literature, we abstract from it assuming that the utility function is homothetic in consumption and (weakly) separable in consumption and leisure (so that $\dot{H}_c = 0$). As for factor 2), it is sufficient to note that the difference between the social weight on different generations and their actual demographic weight constitutes an additional reason for capital taxation, though stemming from equity rather than efficiency considerations. Hence, we will pose our attention on the last two factors and, in particular, on the fourth, which is a new independent source of taxation stemming from our framework.

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 $\mu_g = \mu_p$, $\delta = 0$ and $\gamma_g = \beta$.*

Proof. The proof is straightforward by inspection of eq. (21), which, if the equality $\mu_g = \mu_p$ is satisfied, so that $\frac{\dot{\mu}_g}{\mu_p} = -\delta$, becomes:

$$\tau^k = \frac{1}{f_k + \delta_r} \left[\frac{(\gamma_g - \beta) + \delta \bar{\lambda}(1 + H_c)}{1 + \bar{\lambda}(1 + H_c)} \right];$$

note that this expression, if $\gamma_g = \beta$, is zero only if the probability of death, δ , is zero. Note also that, if $\mu_g = \mu_p$ (and $\dot{H}_c = 0$), then optimal taxes will also be constant through age. ■

¹⁶Moreover, 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 $t > 0$.

Preliminarily, it is worth recalling that, when the conditions above apply (and given that $\dot{H}_c = 0$), the economy mimics the behavior of an ILRA one, so that the zero tax result applies along the transition path. Moreover, it is easy to show that the zero capital income tax rule would apply also in the absence of overlapping generations (i.e. when $\alpha = 0$). On the other hand, even purging out factors 1) to 3), it turns out that the combination of the OLG mechanism and limited life horizons are able to invalidate the Chamley-Judd rule.

More precisely, this source of taxation stems from the fact that individuals, when maximizing their utility, do not take into account the dynamics of the economy generated by the demographic evolution; on the contrary, the government recognizes that when individuals die, leave away a stock of wealth which would be optimal to be confiscated. Thus, the policymaker aims at exploiting this opportunity by raising a corrective capital income tax which is in fact proportional to the probability of death. We will come back on this point after presenting the results at the steady state.

The second best taxation policy, along the steady state path, can be summarized in the following proposition:

Proposition 3 *If the economy converges to a steady state, at such steady state the capital income tax is different from zero unless a) $\mu_g = \mu_p$, $\delta = 0$ and $\gamma_g = \beta$ or b) $\gamma_g > \beta$ and $\delta = 0$.*

Proof. 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 $\bar{\lambda} \rightarrow \lambda$, so that $\tau^k = \frac{\delta}{f_k + \delta_r} \left\{ \frac{-\frac{\dot{\mu}_g}{\delta \mu_p} + \bar{\lambda}(1+H_c)}{\frac{\mu_g}{\mu_p} + \bar{\lambda}(1+H_c)} \right\}$;
moreover, in case $\mu_g = \mu_p$, τ^k is positive and equal to $\frac{\delta}{f_k + \delta_r}$.
2. $\gamma_g > \beta + \delta$. In this case $\bar{\lambda} \rightarrow \infty$, and, again, $\tau^k = \frac{\delta}{f_k + \delta_r}$, (provided that $\dot{\mu}_g$ does not tend to infinity).
3. $\gamma_g < \beta + \delta$. $\bar{\lambda} \rightarrow 0$ and $\tau^k = \frac{1}{f_k + \delta_r} \left[(\gamma_g - (\beta + \delta)) - \frac{\dot{\mu}_g}{\mu_p} \right]$. Moreover, if $\mu_g = \mu_p$, $\tau^k = \frac{1}{f_k + \delta_r} (\gamma_g - \beta)$, which is zero if $\gamma_g = \beta$, otherwise it can be either positive or negative, depending on the relationship between γ_g and β .

■

Again, the economic intuition behind these results can be grasped by reckoning that, when $\gamma_g \geq \beta + \delta$, the tax rate on capital income is proportional to δ , which is the proportion of individuals of each cohort dying at each date t ; now, if we consider that, in the presence of fair life insurance contracts, the individual after tax capital income would be equal to $r(t) a(s, t)$, it follows that, as mentioned above, the optimal tax policy would replicate, at least in the aggregate, the effects of confiscating wealth upon death (which would raise a total revenue amounting to $\delta A(t)$)¹⁷. Finally, 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 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 ($\gamma_g - \beta$). As a consequence, the zero tax result emerges as a very special case (i.e. when $\gamma_g = \beta$).

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 the Chamley Judd rule comes out to be a special case, in 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

¹⁷In other words the optimal tax must leave the individual indifferent between stipulating a fair life insurance contract and paying τ_k in each period, or leaving the whole wealth to the State upon death without any insurance provision and capital income taxation.

Note that if $\delta_r < \delta$ the second scenario would leave individuals better off.

Finally, in case life insurance payments are tax exempt, the same reasoning applies; the only change is in the denominator of the tax expression, which would be simply f_k (instead of $f_k + \delta_r$).

the weight the government attaches to each generation and its actual demographic size; c) the difference between the government and individual intertemporal discount rates; d) the probability of death, δ (inducing a limit to individual lifetime).

The first factor has been widely discussed in the literature, while the second descends from intuitive equity arguments. The economic intuition underlying the role of the remaining two factors 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; however, even if the two intertemporal rates do coincide, it is optimal to set a positive tax rate proportional to δ , because this would mimic the effects of confiscating individual wealth upon death in an economy without life insurance.

Finally, from the analysis above it turns out that the violation of the Chamley Judd rule does crucially depend upon the assumption of overlapping generations, combined with that of finite horizons. In fact both 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.

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] L.E. Jones, R.E. Manuelli, P.E. Rossi, On the Optimal Taxation of Capital Income, *J. Econ. Theory* **73** (1997), 93-117.
- [13] K.L. Judd, Redistributive Taxation in a Simple Perfect Foresight Model, *J. Public Econ.* **28** (1985), 59-83.
- [14] K.L. Judd, Optimal Taxation and Spending in General Competitive Growth Models, *J. Public Econ.* **71** (1999), 1-26.
- [15] 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.
- [16] F.P. Ramsey, A contribution to the theory of taxation, *Econ. J.* **37** (1927), 47-61.
- [17] T. Renström, Optimal Dynamic Taxation, in: S.B. Dahiya (Ed.), *The Current State of Economic Science*, Vol. 4, Spellbound Publications, 1999, pp1717-1743.
- [18] P. Weil, Overlapping families of infinitely-lived agents, *J. Public Econ.* **38** (1989), 183-198.
- [19] I. Yakadina, Optimal Capital-Labor Taxes under Uncertainty and Limits on Debt, *Universitat Pompeu Fabra*, mimeo, 2001.

- [20] 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: $\tilde{w}(s, t) = -\frac{U_{l(s,t)}}{U_{c(s,t)}}$, $[\tilde{r}(s, t) + \tilde{\delta}_r(s, t)] = \left(\beta + \delta - \frac{\dot{p}(s,t)}{p(s,t)}\right)$, 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:

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

then, by exploiting the expression for $\dot{p}(s, t)$ we get¹⁸:

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

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

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

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

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:

¹⁸The equations below hold $\forall s$.

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

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

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

which is eq. (12) in the text. ■

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 [15], suppose the problem has the form

$$\max \int \int F(t, s, x(t, s), x_t(t, s), x_s(t, s)) ds dt$$

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:

$$\int \int q(t, s, x(t, s), x_t(t, s), x_s(t, s)) ds dt = 0,$$

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

$$\int \int (F + \lambda q) ds dt.$$