

Optimal income taxation and migration

Valeria De Bonis¹  · Luca Spataro² 

Published online: 5 February 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract We tackle the issue of optimal dynamic taxation of capital income in an economy with disconnection as in Weil (J Public Econ 38:183–198, 1989), generated by migration and intra-family altruism. We show that, when the government aims at correcting such a disconnection using time-varying weights in the social welfare function, then there is room for nonzero capital income taxation, both in the short and in the long run.

Keywords Optimal dynamic taxation · Migration · Social discounting

JEL Classification E62 · H21

1 Introduction

Starting from the seminal works by Judd (1985) and Chamley (1986), the issue of dynamic optimal capital income taxation has been analysed by a number of researchers. In particular, Judd (1999) has shown that the zero-tax rate result stems 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 (general equilibrium) elasticity of consumption is

✉ Luca Spataro
luca.spataro@uniipi.it

Valeria De Bonis
valeria.debonis@uniroma1.it

¹ Istituto di Economia e Finanza, Sezione del Dipartimento di Studi Giuridici, Filosofici ed Economici, Facoltà di Giurisprudenza, Università di Roma “La Sapienza”, Rome, Italy

² Dipartimento di Economia e Management, Università di Pisa, Pisa, Italy

constant over time. In infinitely lived representative agent (ILRA) models,¹ along the transition path, this condition holds only if the utility function is assumed (weakly) separable in consumption and leisure and homothetic in consumption. It is, instead, necessarily satisfied in the long run, thus leading to the well-known result of optimal zero capital income taxes.

The result appears at odds with reality, since the existence of positive capital income taxes is often observed. Several theoretical models have tried to find a rationale for capital income taxation.

One avenue has been allowing the government's discount rate to differ from the households' one. De Bonis and Spataro (2005) find that the zero capital tax result does not generally hold when the government discount rate differs from the individual one, due to "Pigouvian" arguments²: along the transition path, even with separable and homothetic utility functions, capital income is taxed (subsidized) when the government is less (more) impatient than individuals are. Instead, they obtain an asymmetry as for the steady state, since in the long run only the case for subsidizing capital is confirmed. This is because, when the government discounts the future more heavily than private agents do, the explosive distortionary effect of taxation impedes to hit capital income.³ Reis (2012) finds that capital income taxes are positive even in the long run if one introduces lack of commitment for the government alongside the difference in the discount rates: if the government is more impatient than households are, it will increase debt, but the lack of commitment will force it to save to make future policies credible. This will prevent the government from becoming too indebted, which would force it to resort only to the more efficient labour taxes.⁴

Another avenue to obtain the result of the violation of the zero tax rule also in the long run is abandoning the standard ILRA framework in favour of overlapping generations models with life cycle (OLG-LC).⁵ This outcome can be understood by reckoning that in such a set-up optimal consumption and labour, or, more precisely, the (general equilibrium) elasticity of consumption, are generally not constant over life and even at the steady state, due to life-cycle behaviour.

¹ See Atkeson et al. (1999) and Chari et al. (1999).

² The difference between the government and individuals' discount rates can be explained in terms of myopia of the latter, or of political instability, where a government takes into account that it might lose office and therefore values the future less than individuals do (see Grossman and Van Huyck 1988).

³ By the same reasoning, the asymmetry disappears in the special case of logarithmic utility function, since the anticipated policy path does not affect current individual choices and, thus, the cumulative distortionary effect of taxes is ruled out (see also Lansing 1999 and Reinhorn 2013). For a recent discussion on nonzero taxation of capital income, see also Straub and Werning (2014).

⁴ Another extension of the basic framework aimed at obtaining a positive taxation of capital consists in the introduction of uncertainty. However, even if Zhu (1992) points out that in such a case the capital income tax may be nonzero in the long run, Chari et al. (1994) show via simulations that the average value of the optimal capital income tax is very close to zero. Other authors have explored the role of market incompleteness in stochastic environments: Chari and Kehoe (1999), for example, consider an economy with state contingent returns on debt acting as shock absorbers. They show that if the government cannot issue bonds with state contingent returns, capital income taxes can be chosen to overcome this problem. More recently, on the role of idiosyncratic uncertainty and private information in invalidating the Chamley–Judd result, see Golosov et al. (2003), Kocherlakota (2005) and Albanesi and Sleet (2006).

⁵ See Atkinson and Sandmo (1980) and Erosa and Gervais (2002).

Both ILRA and OLG-LC models are characterized by very simple population dynamics. In this work, we extend the previous analysis on optimal taxation to a model with a “disconnected”⁶ economy, generated by migration and limited altruism. More precisely, we consider a model with individuals who are altruistic only towards their own descendants and disregard the well-being of immigrants.⁷ Hence, the entrance of new dynasties in each period prevents the economy from behaving as a single representative individual one and creates an “overlapping dynasties” mechanism.

The literature has widely analysed the interconnections between migration and fiscal policies, in particular those stemming from differences between natives and immigrants.⁸ In this paper, however, we only consider the different time of entry into the economy as a heterogeneity dimension. This is because the paper aims at analysing the effects of the disconnectedness caused by migration on capital income taxation once the policymaker, differently from residents, aims at correcting it by taking the utility of new entrants properly into account in the social welfare function. The issue of the weight to be assigned to future generations has long been debated and addressed in recent works.⁹ However, so far the literature has assumed a constant social discount rate (see, for instance, Erosa and Gervais 2002). We depart from the traditional approach by assuming that the social discount rate is time-varying and, precisely, inversely correlated with the age of the dynasty.¹⁰ In particular, we consider the case in which such weights coincide with the actual demographic weights of dynasties, which, due to immigration flows, are declining through time.¹¹ In fact, such an assumption, which seems reasonable and consistent with a Benthamite approach, turns out to be equivalent to assuming a constant intergenerational discount rate and a social intertemporal discount rate that differs from the individual one.

In this scenario, the main result we find is that the capital income tax can be nonzero, even in the absence of life-cycle behaviour. The result directly stems from Pigouvian arguments: declining dynasty’s weights lead the government to correct private accumulation of capital. More precisely, since the government cares less about the future well-being of the dynasty than the dynasty does, it wishes to tax future consumption more than present one, which implies a positive capital income tax.

⁶ The relevance of the “disconnectedness” of the economy has been firstly analysed by Weil (1989) in the context of the validity of the Ricardian equivalence proposition.

⁷ For a presentation of this model in continuous time, see Barro and Sala-i-Martin (1999), chapter 9.

⁸ See, for instance, Razin et al. (2002) and Occhino (2008) on fiscal incentives and the relationship between migration and welfare policies; Solé-Auró and Crimmins (2008) on the differences in consumption patterns of natives and immigrants; Czaika and Parsons (2017) for a recent contribution on the effects of international policy agreements on the characteristics of immigrants; and Bell and Eiser (2016) for the effects of the interplay of migration and fiscal policies on the performance of labour markets.

⁹ For example, Ramsey (1928) considered constant discounting unacceptable from an ethical standpoint; see, among others, Bernheim (1989), Michel (1990) and Caplin and Leahy (2004).

¹⁰ A notable exception is the work by Farhi and Werning (2005), who obtain a time-varying social discount rate by assuming that the government values future generations directly and not simply through the altruism of the current generation and in the absence of an OLG framework. See also Spataro and De Bonis (2008) for an analysis of the issue in the context of a perpetual youth economy.

¹¹ For a general analysis of the taxation of savings and, in particular, inheritance taxes in a similar framework, see De Bonis and Spataro (2010).

Even if the rationale is the same as in the ILRA framework in De Bonis and Spataro (2005), disconnection in the economy generates positive capital income taxes also in the long run. Differently from the result obtained in Reis (2012), lack of commitment is not necessary to obtain the outcome.

The work proceeds as follows: in the section that follows, 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.

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 and have infinite lives, differ as for their date of entry into the economy, s ; natives are supposed to have entered the economy at time $s = -1$, while migrants start entering at time 0 at a given rate α ; both types of individuals have a constant rate of growth n : as a consequence, the population growth rate is equal to $\gamma \equiv (1 + \alpha)(1 + n) - 1$; the whole population at time t has cardinality:

$$N_t = (1 + \gamma)^{t+1},$$

with the size of population at time -1 normalized to 1; the size of each dynasty (started by the entry of the founder) is¹²:

$$P_{s,t} = \alpha^d (1 + \alpha)^{d \cdot s} \cdot (1 + n)^{t+1}; t \geq 0, -1 \leq s \leq t$$

with $d = 0$ if $s = -1$, 1 otherwise. Moreover, all individuals offer labour 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 factors into production–consumption units. Finally, the government can finance an exogenous stream of public expenditure G_t by issuing internal debt B_t and by raising proportional taxes both on interests and wages, referred to as $\tau_{s,t}^k$ and $\tau_{s,t}^l$, respectively. Notice that taxes can be conditioned on the date of birth of dynasties.

¹² To see this, consider period 0: the dynasty of the natives consists of 1 native (the founder, born in period -1) and n children, so that $P_{-1,0} = (1 + n)$ and the dynasties of the first immigrants are formed by $P_{0,0} = \alpha(1 + n)$ individuals, with $N_0 = (1 + \alpha)(1 + n)$. In period 1, the dynasty of natives is formed by the founder, n children and $n(1 + n)$ nephews, such that $P_{-1,1} = (1 + n)^2$. In the same period 1, the dynasties of the first immigrants are formed by $\alpha(1 + n)$ founders and $\alpha(1 + n)n$ children, so that $P_{0,1} = \alpha(1 + n)^2$, while new immigrants have entered the economy, equal to a fraction α of the previous period population and their current children, i.e. $P_{1,1} = \alpha(1 + \alpha)(1 + n)^2$, so that $N_1 = (1 + \alpha)^2(1 + n)^2$. Generalizing, we end up with the above expressions for $P_{s,t}$ and N_t .

2.1 Private agents

Agents’ preferences can be represented by the following instantaneous utility function:

$$U(c_{s,t}, l_{s,t}),$$

where $c_{s,t}$ and $l_{s,t}$ are instantaneous consumption and labour supply, respectively, of the dynasty founded in period s , as of instant t . Such a utility function is strictly increasing in consumption and decreasing in labour, strictly concave, and satisfies the standard Inada conditions. Since we assume that individuals care about the well-being of their children, agents maximize the following utility function:

$$\max_{\{c_{s,t}, l_{s,t}\}_s^\infty} \sum_{t=s}^\infty \left(\frac{1+n}{1+\beta}\right)^{t-s} U(c_{s,t}, l_{s,t}) \tag{1}$$

sub:

$$a_{s,t} = \frac{(1 + \tilde{r}_{s,t})}{(1 + n)} a_{s,t-1} + \tilde{w}_{s,t} l_{s,t} - c_{s,t},$$

$$\lim_{t \rightarrow \infty} a_{s,t} \frac{(1+n)^{t-s}}{\prod_{i=s+1}^t (1 + \tilde{r}_{s,i})} = 0, \quad a_{s,s-1} = 0, \quad a_{0,-1} = \bar{k}_{-1}, \tag{2}$$

where β is the intertemporal discount rate, with $1 > \beta > n > 0$, a is the agent’s wealth, while $\tilde{w}_{s,t} = w_t(1 - \tau_{s,t}^l)$ and $\tilde{r}_{s,i} = r_t(1 - \tau_{s,t}^k)$ are the net-of-tax factor prices.

The FOCs of this problem imply:

$$\left(\frac{1+n}{1+\beta}\right)^{t-s} U_{c_{s,t}} = p_{s,t} \tag{3}$$

$$\left(\frac{1+n}{1+\beta}\right)^{t-s} U_{l_{s,t}} = -p_{s,t} \tilde{w}_{s,t} \tag{4}$$

$$\frac{(1 + \tilde{r}_{s,t+1})}{(1 + n)} p_{s,t+1} = p_{s,t}, \tag{5}$$

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 shadow price of wealth. These conditions yield:

$$\frac{U_{c_{s,t}}}{U_{c_{s,t+1}}} = \frac{(1 + \tilde{r}_{s,t+1})}{1 + \beta} \tag{6}$$

$$\frac{U_{l_{s,t}}}{U_{c_{s,t}}} = -\tilde{w}_{s,t}. \tag{7}$$

which provide the usual relations between individual’s marginal rate of substitution and prices along an optimal trajectory.

2.2 Firms

We assume that firms are identical and operate in a competitive environment; they hire capital and labour services according to their market prices (gross of taxes). This means that, for each firm i , profit maximization yields:

$$\frac{dF(K_{t-1}^i, L_t^i)}{dK_{t-1}^i} = r_t \tag{8}$$

$$\frac{dF(K_{t-1}^i, L_t^i)}{dL_t^i} = w_t. \tag{9}$$

Note that capital is assumed to enter the production process with a one period lag. Assuming a CRS technology, such conditions can also be expressed for the economy as a whole, in per capita terms:

$$f_{k_{t-1}} = \frac{r_t}{1 + \gamma} \tag{8’}$$

$$f_{l_t} = w_t. \tag{9’}$$

where $l_t \equiv \frac{L_t}{N_t} = \sum_{s=0}^t v_{s,t} l_{s,t}$ and $k_t \equiv \frac{K_t}{N_t} = \sum_{s=0}^t v_{s,t} k_{s,t}$, with $v_{s,t} \equiv \frac{P_{s,t}}{N_t} = \frac{\alpha^d}{(1+\alpha)^{t+1-d \cdot s}}$ and $d = 0$ if $s = -1, 1$ otherwise, the weight of dynasty s in the economy population at period t .

2.3 The government and market clearing conditions

The government is assumed to finance an amount of exogenous public expenditure by levying taxes on capital and labour income and by issuing debt in the absence of lump sum taxation. In order to rule out the problem of time inconsistency, we suppose that the government has access to a commitment technology that ties it to the announced path of distortionary tax rates whenever the possibility of lump sum taxation arises. The only constraints on the possibility of debt issuing are the usual no-Ponzi game condition and the initial condition $B_0 = \bar{B}$. Thus, one obtains the usual equation for the dynamics of aggregate debt:

$$B_t = (1 + r_t)B_{t-1} + G_t - T_t, \tag{10}$$

where $T_t = \sum_{s=0}^t P_{s,t} [\tau_{s,t}^l w_t l_{s,t} + \frac{\tau_{s,t}^k r_t a_{s,t-1}}{(1+n)}]$, Eq. (10) can also be written, in per capita terms:

$$b_t = \left(\frac{1 + r_t}{1 + \gamma} \right) b_{t-1} + g_t - \tau_t. \tag{11}$$

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

$$A_t = K_{t-1} + B_t. \tag{12}$$

3 The Ramsey problem

Since we adopt the primal approach to the Ramsey (1927) problem,¹³ a key point is restricting the set of allocations among which the benevolent government can choose to those that can be decentralized as a competitive equilibrium.¹⁴ Thus, in this paragraph we define a competitive equilibrium and the constraints that must be imposed on the policymaker problem in order to achieve such a competitive outcome.

The first constraint is the implementability constraint, i.e. the dynasty budget constraint with prices substituted for by exploiting the individual FOCs (for the derivation see Appendix):

$$\sum_{t=s}^{\infty} \left(\frac{1 + n}{1 + \beta} \right)^{t-s} (U_{c_{s,t}} c_{s,t} + U_{l_{s,t}} l_{s,t}) = 0, \forall s \tag{13}$$

which is referred to as the “implementability constraint”.¹⁵

As for the second constraint, summing Eq. (2) over population to get aggregate wealth, subtracting Eq. (10) and exploiting the market clearing condition, we get:

$$y_t \geq c_t + k_t - \frac{k_{t-1}}{1 + \gamma} + g_t, \tag{14}$$

where y is output in per worker terms. Such an expression is usually referred to as the “feasibility constraint” (see Appendix for a formal derivation).

We can now give the following definition, supposing that the policy is introduced in period 0:

Definition 1 A competitive equilibrium is: a) an infinite sequence of policies $\pi = \{\tau_{s,t}^k, \tau_{s,t}^l, b_t\}_{t=0}^{\infty}$, b) allocations $\{c_{s,t}, l_{s,t}, k_t\}_{t=0}^{\infty}$ and c) prices $\{w_t, r_t\}_{t=0}^{\infty}$ such that, at each instant t : b) satisfies Eq. (1) subject to (2), given a) and c); c) satisfies Eq. (8’) and (9’); Eqs. (14) and (11) are satisfied.

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

¹³ See Stiglitz (2015) for a review.

¹⁴ See Atkinson and Stiglitz (1980); the alternative “dual” approach takes prices and tax rates as control variables (see, for instance, Chamley 1986).

¹⁵ As for the first dynasty, the implementability constraint takes the form:
 $\sum_{t=s}^{\infty} \left(\frac{1+n}{1+\beta} \right)^{t-s} (U_{c_{s,t}} c_{s,t} + U_{l_{s,t}} l_{s,t}) = U_{c_{0,0}}(1 + \tilde{r}_{0,0})\bar{k}_{-1}$.

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 proof of the reverse (any allocation satisfying implementability and feasibility is a competitive equilibrium) is provided in Appendix. □

3.1 Solution

We now assume that the policymaker maximizes a utilitarian social welfare function, which is a weighted sum of the dynasties’ utilities, subject to the constraints presented above.¹⁶ By defining the auxiliary function:

$$W_s = \sum_{t=s}^{\infty} \left(\frac{1+n}{1+\beta} \right)^{t-s} [\mu_{s,t} U(c_{s,t}, l_{s,t}) + \lambda_s (U_{c_{s,t}} c_{s,t} + U_{l_{s,t}} l_{s,t})],$$

the policymaker’s problem is the following:

$$\max_{\{c_{s,t}, l_{s,t}, k_t\}_0^{\infty}} \sum_{s=0}^{\infty} W_s$$

sub:

$$y_t = c_t + k_t - \frac{k_{t-1}}{1+\gamma} + g_t, \quad \forall t$$

$$\lim_{t \rightarrow \infty} \frac{k_t}{\prod_{i=1}^t (1+f_{k_i})} = 0, \quad k_{-1} = \bar{k}.$$

where $\mu_{s,t}$ is the weight that the government attaches to dynasty s .¹⁷ Note that, differently from previous works, $\mu_{s,t}$ is allowed to vary with time.

The FOCs of the problem with respect to consumption and capital are, respectively:

$$\frac{\partial W_t}{\partial c_{s,t}} = 0 \Rightarrow \left(\frac{1+n}{1+\beta} \right)^{t-s} U_{c_{s,t}} [\mu_{s,t} + \lambda_s (1 + H_{c_{s,t}})] = \phi_t v_{s,t} \tag{15}$$

$$\frac{\partial W_t}{\partial k_t} = 0 \Rightarrow \phi_t = \left(\frac{1}{1+\gamma} + f_{k_t} \right) \phi_{t+1} \tag{16}$$

¹⁶ In our model, thus, the reference point in the government maximization problem is the utility of the dynasties. The dynamic dimension of income has been recently put forward in the context of the distributional effects of taxation (Piketty 2014; Atkinson 2015; Kanbur and Stiglitz 2015; Halvorsen and Thoresen 2017). Even if our framework is very simple, since dynasties only differ as for their date of entry into the economy and we adopt an utilitarian welfare function, we contribute to this literature by adding efficiency aspects to the analysis of the taxation of dynastic income.

¹⁷ We omit the government budget constraint since, by Walras’ law, it is satisfied if the implementability and feasibility constraints hold.

where λ_s is the multiplier associated with the implementability constraint, $H_{c_{s,t}} = \frac{U_{cc_{s,t}}c_{s,t} + U_{lc_{s,t}}l_{s,t}}{U_{c_{s,t}}}$, which is usually referred to as the “general equilibrium elasticity of consumption,” and ϕ_t is the multiplier associated with the feasibility constraint.

By taking the first order condition relative to consumption as for period $t + 1$ and dividing Eq. (15) by it, we get¹⁸:

$$\frac{\left(\frac{1+n}{1+\beta}\right)^{t-s} U_c[\mu + \lambda(1 + H_c)]}{\left(\frac{1+n}{1+\beta}\right)^{t+1-s} U_c^{+1} [\mu^{+1} + \lambda(1 + H_c^{+1})]} = \frac{\phi_t v}{\phi_{t+1} v^{+1}} \tag{17}$$

and, by exploiting Eqs. (6), (8') and (16) and by reckoning that $\frac{v}{v^{+1}} = (1 + \alpha)$, we get:

$$\frac{1 + \tilde{r}^{+1}}{1 + r^{+1}} = \frac{\mu^{+1} + \lambda(1 + H_c^{+1})}{\mu + \lambda(1 + H_c)} \tag{18}$$

which provides the implicit expression for the optimal capital income tax.

4 Discussion of the results

We now discuss the results concerning capital income taxation at the steady state.

By inspection of the tax equation, it emerges that there are two independent forces determining the level of τ^k : (1) the dynamics of H_c and (2) the dynamics of the social intergenerational weight, which is new.

We can now state the following proposition:

Proposition 2 *If the economy converges to a steady state, along the transition path, for $t > t_0$, and at the steady state the tax on capital income is different from zero unless $\frac{\mu^{+1} + \lambda(1 + H_c^{+1})}{\mu + \lambda(1 + H_c)} = 1$.*

Proof The proof is straightforward by inspection of Eq. (18). □

Factor 1) has been widely discussed in the literature: $H_c = H_c^{+1}$ obtains, for example, by assuming homotheticity in consumption and (weak) separability in consumption and leisure in the utility function. 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 H_c can vary with age even at the steady state. However, as shown in Eq. (18), even in the absence of a life cycle, in the present model the nonzero tax rule can still apply.

¹⁸ From now onward, we omit the s and t indicators, whenever this does not cause ambiguity: hence, notation X^{+1} stands for $X_{s,t+1}$.

The role of factor 2) can be isolated by supposing $H_c = H_c^{+1}$.¹⁹ Then, $\frac{1+\tilde{r}^{+1}}{1+r^{+1}} = \frac{\mu^{+1}+\lambda(1+H_c)}{\mu+\lambda(1+H_c)}$. As for the choice of the social weight, we consider two exemplifying cases, chosen to allow the government to distinguish either on the basis of the dynasty only (only the index s of the weight μ varies), or of both the dynasty and its age (both s and t vary).

- (a) $\mu_{s,t} = \mu$. The weight assigned to each dynasty by the government is constant through time. Then, $\frac{1+\tilde{r}^{+1}}{1+r^{+1}} = 1$, which implies the absence of capital income taxation. The assumption of a constant μ is made in the existing OLG-LC models (see, in particular, Erosa and Gervais 2002 and Garriga 2003).²⁰
- (b) $\mu_{s,t} \neq \mu_{s,t+1}$. The weight assigned to the dynasty varies through time, that is, through the age of the dynasty. In this case, the capital income tax is different from zero. Let us consider the Benthamite case in which the social weight of each dynasty is equal to its actual demographic weight within the population, i.e. $\mu_{s,t} = \nu_{s,t}$. Given our assumption of a constant rate of migration arrivals α , the relative size of each dynasty is decreasing through time, so that $\mu^{+1}/\mu = 1/(1+\alpha)$ for each dynasty. Hence, the tax rate is, in general, positive. This result stems from the fact that the government cares less about the future utilities of the dynasty than the dynasty does and therefore discriminates future consumption in favour of the present one. Given that the dynasty's weight decreases with time, also the tax on each dynasty's capital income will decrease with age and will tend to zero for $t - s \rightarrow \infty$, that is, for the oldest dynasties (since their social weight tends to zero).

For illustrative purposes, in Fig. 1 we report the dynamics of the capital income tax for a generic dynasty born in year s , endowed with a CES utility function with separable consumption and leisure. According to this example, in which the government uses demographic weights as social weights, the capital income tax would halve in fifty years and halve again in the next fifty years, then tending asymptotically to zero. Notice that the same figure can be interpreted as the steady-state tax structure as a function of the age (i.e. size) of each dynasty. We again notice that the capital income tax will be nonzero and tending asymptotically towards zero as age increases.

Another way to grasp the forces delivering this result is to note that the presence of a social weight declining over time corresponds to a government time discounting that differs from the private one. This can be seen by rewriting the problem of the policymaker as follows:

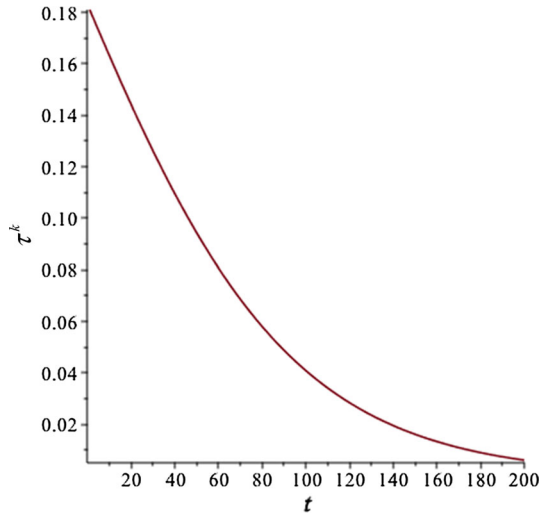
$$\max_{\{c_{s,t}, l_{s,t}, k_t\}_0^\infty} \sum_{s=0}^\infty \hat{\mu}_s \hat{W}_s$$

¹⁹ This case occurs, for instance, when the utility function is of the form: $U = \frac{c^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} + V(l)$, where

$$H_c = -\frac{1}{\sigma}.$$

²⁰ Generally speaking, differences in weight can be connected to different characteristics of natives and immigrants, and of different generations of immigrants; this would be of relevance for the choice of taxation instruments other than capital income taxation.

Fig. 1 Dynamics of the capital income tax for a generic dynasty. Parameters specification: $s = 1, \sigma = 2, \lambda = 0.05, \alpha = 2\%, r = 5\%$



where $\hat{\mu}_s = \alpha^d(1 + \alpha)^s$ and

$$\hat{W}_s = \sum_{t=s}^{\infty} \left(\frac{1+n}{1+\beta} \right)^{t-s} \left[(1 + \alpha)^{-t} U_{s,t} + \frac{(1 + \alpha)^{-s}}{\alpha^d} \lambda_s (U_{c_{s,t}} c_{s,t} + U_{l_{s,t}} l_{s,t}) \right].$$

Hence, our result can be interpreted in terms of Pigouvian correction: since the government is less patient than individuals are, it is optimal to hit capital income in order to slow down the accumulation of private capital and, thus, increase present consumption. Finally, note that our result would still apply if lump sum taxes were available. In fact, the result would hold even if the implementability constraint did not bind, i.e. for the case of $\lambda = 0$, i.e. if lump sum taxes were available.

Thus, the case of constant dynasty weights corresponds to a social discount rate that is equal to the private one; consequently, if the elasticity of consumption is constant, there is no reason for correcting individual choices.

In the case of time-varying dynasty weights, instead, if these are chosen based on the dynasty relative size, the social intertemporal discount rate is smaller than the private one. Therefore, the government discriminates against future consumption.

Finally, we can provide the shape of the labour income tax under sufficiently general assumptions on the utility function (that is, weak complementarity between consumption and leisure):

Proposition 3 *Suppose that the utility function is such that $U_{c,l} \leq 0$; if the economy converges to a steady state, then the optimal labour tax is positive.*

Proof Taking the first derivative with respect to labour supply yields:

$$\frac{\partial W_t}{\partial l_{s,t}} = 0 \Rightarrow \left(\frac{1+n}{1+\beta} \right)^{t-s} U_{l_{s,t}} [\mu_{s,t} + \lambda_s(1 + H_{l_{s,t}})] = -\phi_t v_{s,t} f_{l_t} \quad (19)$$

where $H_{l_{s,t}} = \frac{U_{ll_{s,t}}l_{s,t} + U_{lc_{s,t}}c_{s,t}}{U_{l_{s,t}}}$ is “general equilibrium elasticity of labour.” Combining (19) with (15), exploiting (7) and (9') and collecting terms it descends that:

$$\tau_{s,t}^l = \lambda_s \frac{H_{l_{s,t}} - H_{c_{s,t}}}{\mu_{s,t} + \lambda_s(1 + H_c)} \tag{20}$$

By (19), the denominator of (20) is positive and so is λ_s . Exploiting the definitions of H_l and H_c , it can be seen that a sufficient condition for the numerator of (20) to be positive is $U_{c,l} \leq 0$. □

5 Conclusions

We reconsider the issue of optimal capital income taxation in an OLG economy with dynastic altruism and migration applying the primal approach to the Ramsey problem.

The thrust of the paper is that the Chamley–Judd zero tax rule comes out not to apply if the policymaker wishes to correct the disconnection of the economy by attaching to each dynasty a social weight that varies with time. Moreover, if the latter corresponds to each dynasty’s actual share within the population the optimal rule prescribes positive taxation of capital income, even in the absence of life-cycle behaviour, and positive taxation of labour income as well.

Taking into account the dynamics of population in the social welfare function, while being absent in both “representative agent” and traditional OLG models, besides being interesting per se, seems appropriate for models in which population “matters” and can represent a useful criterion for normative prescriptions concerning the real world, in which the dynamics of the population are typically very complicated. In particular, our model might be extended by connecting the date of entry into the economy to some population characteristics that are relevant to taxation policy (heterogeneity in tastes, skills, income and wealth levels).

Compliance with ethical standards

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest The authors declare that they have no conflict of interest.

6 Appendix

6.1 Derivation of the implementability constraint

In order to obtain the implementability constraint, write Eq. (2) in its intertemporal form:

$$\sum_{t=s}^{\infty} \frac{(c_{s,t} - \tilde{w}_{s,t}l_{s,t})}{\prod_{i=s+1}^t (1 + \tilde{r}_{s,i})} = 0. \tag{21}$$

Since:

$$\frac{U_{c_{s,t}}}{U_{c_{s,t+1}}} \frac{(1 + \beta)}{(1 + n)} = \frac{p_{s,t}}{p_{s,t+1}} = \frac{(1 + \tilde{r}_{t+1})}{(1 + n)},$$

we have

$$\frac{1}{(1 + n)^{t-s}} \prod_{i=s+1}^t (1 + \tilde{r}_{s,i}) = \frac{p_{s,s}}{p_{s,s+1}} \frac{p_{s,s+1}}{p_{s,s+2}} \dots \frac{p_{s,t-1}}{p_{s,t}}.$$

By substituting into Eq. (21), we obtain

$$\sum_{t=s}^{\infty} \frac{(c_{s,t} - \tilde{w}_{s,t} l_{s,t}) p_{s,t}}{p_{s,s}} = 0$$

and exploiting the FOCs from the individual maximization problem, we get

$$\sum_{t=s}^{\infty} \left(\frac{1 + n}{1 + \beta} \right)^{t-s} (U_{c_{s,t}} c_{s,t} + U_{l_{s,t}} l_{s,t}) = 0.$$

which is Eq. (13) in the text.

6.2 Derivation of the feasibility constraint

To derive the feasibility constraint, first aggregate Eq. (2) over population at time t (notice that $a_{s,s-1} = 0$):

$$\sum_{s=0}^t P_{s,t} a_{s,t} = \sum_{s=0}^t P_{s,t} \left[\frac{(1 + r_t)}{(1 + n)} a_{s,t-1} + w_t l_{s,t} - c_{s,t} - \frac{\tau_{s,t}^k r_t a_{s,t-1}}{1 + n} - \tau_{s,t}^l w_t l_{s,t} \right]. \tag{22}$$

and by recalling that $A_t \equiv \sum_{s=0}^t P_{s,t} a_{s,t}$, $P_{s,t} = P_{s,t-1}(1 + n)$ and $P_{t,t-1} = 0$, so that

$\sum_{s=0}^t P_{s,t-1} a_{s,t-1} = A_{t-1}$ we can rewrite Eq. (22) as follows

$$A_t = (1 + r_t)A_{t-1} + w_t L_t - C_t - T_t$$

where C_t is aggregate consumption. Finally, by subtracting Eq. (10) and exploiting the market clearing condition we obtain

$$K_t = (1 + r_t)K_{t-1} + w_t L_t - C_t - G_t$$

which, in per capita terms, becomes

$$k_t = \frac{(1 + r_t)}{(1 + \gamma)} k_{t-1} + w_t l_t - c_t - g_t$$

where $\frac{r_t k_{t-1}}{(1+\gamma)} + w_t l_t \equiv y_t$ due to CRS.

6.3 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: $\tilde{w}_{s,t} = -\frac{U_{l_{s,t}}}{U_{c_{s,t}}}$, $\tilde{r}_{s,t} = \frac{p_{s,t}}{p_{s,t+1}}(1+n) - 1$, with $p_{s,t} = U_{c_{s,t}}(\frac{1+n}{1+\beta})^{t-s}$, $\forall s$ and $\forall t$, and a sequence of before tax prices: $f_{k_{t-1}} = \frac{r_t}{1+\gamma}$, $f_{l_t} = w_t$. Therefore, 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 $U_{c_{s,t}}$ and $U_{l_{s,t}}$ by using the expressions above:

$$\sum_{t=s}^{\infty} (p_{s,t}c_{s,t} - \tilde{w}_{s,t}p_{s,t}l_{s,t}) = 0 \quad \forall s.$$

Then, by recursively using the expression $p_{s,t} = \frac{p_{s,s}(1+n)^{t-s}}{\prod_{i=s+1}^t (1+\tilde{r}_{s,i})}$ we get

$$\sum_{t=s}^{\infty} \frac{(1+n)^{t-s}}{\prod_{i=s+1}^t (1+\tilde{r}_{s,i})} p_{s,s}(c_{s,t} - \tilde{w}_{s,t}l_{s,t}) = 0.$$

Finally, by eliminating $p_{s,s}$ and defining $c_{s,t} - \tilde{w}_{s,t}l_{s,t} = -(1+n)q_{s,t} + (1+\tilde{r}_{s,t})q_{s,t-1}$, we get

$$\sum_{t=s}^{\infty} \frac{(1+n)^{t-s}}{\prod_{i=s+1}^t (1+\tilde{r}_{s,i})} [-(1+n)q_{s,t} + (1+\tilde{r}_{s,t})q_{s,t-1}] = 0$$

that turns out to be:

$$-\lim_{t \rightarrow \infty} \frac{(1+n)^{t-s}}{\prod_{i=s+1}^t (1+\tilde{r}_{s,i})} q_{s,t} + q_{s,s} = 0$$

which holds if $q_{s,t} = a_{s,t}$ and $\lim_{t \rightarrow \infty} \frac{(1+n)^{t-s}}{\prod_{i=s+1}^t (1+\tilde{r}_{s,i})} a_{s,t} = 0$.

As for the public sector budget constraint, by aggregating the individuals' budget constraints over population at time t and expressing them in per capita terms, we get

$$a_t = \frac{(1 + r_t)}{(1 + \gamma)} a_{t-1} + w_t l_t - c_t - \tau_t. \quad (23)$$

Finally, by subtracting the feasibility constraint

$$k_t = \frac{(1 + r_t)}{(1 + \gamma)} k_{t-1} + w_t l_t - c_t - g_t$$

and defining $b_t = -k_{t-1} + a_t$, we obtain

$$b_t = \frac{(1 + r_t)}{(1 + \gamma)} b_{t-1} + g_t - \tau_t$$

which is Eq. (11) in the text. □

References

- Albanesi, S., & Sleet, C. (2006). Dynamic optimal taxation with private information. *Review of Economic Studies*, 73, 1–30.
- Atkeson, A., Chari, V. V., & Kehoe, P. J. (1999). Taxing capital income: A bad idea. *Federal Reserve Bank of Minneapolis Quarterly Review*, 23, 3–17.
- Atkinson, A. B. (2015). *Inequality: What can be done*. Cambridge: Harvard University Press.
- Atkinson, A. B., & Sandmo, A. (1980). Welfare implications of the taxation of savings. *The Economic Journal*, 90, 529–549.
- Atkinson, A. B., & Stiglitz, J. E. (1980). *Lectures on public economics*. London: McGraw-Hill.
- Barro, R. J., & Sala-i-Martin, X. (1999). *Economic growth*. Cambridge: The MIT Press.
- Bell, N. F., & Eiser, D. (2016). Migration and fiscal policy as factors explaining the labour-market resilience of UK regions to the great recession. *Cambridge Journal of Regions, Economy and Society*, 9(1), 197–215.
- Bernheim, B. D. (1989). Intergenerational altruism, dynastic equilibria and social welfare. *Review of Economic Studies*, 56, 119–128.
- Caplin, A., & Leahy, J. (2004). The social discount rate. *Journal of Political Economy*, 112, 1257–1268.
- Chamley, C. (1986). Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica*, 54, 607–622.
- Chari, V. V., Christiano, L. J., & Kehoe, P. J. (1994). Optimal fiscal policy in a business cycle model. *Journal of Political Economy*, 102, 617–652.
- Chari, V. V., & Kehoe, P. J. (1999). Optimal fiscal and monetary policy. In J. B. Taylor & M. Woodford (Eds.), *Handbook of macroeconomics* (Vol. 1, pp. 1670–1745). Amsterdam: North-Holland.
- Czaika, M., & Parsons, C. R. (2017). The gravity of high-skilled migration policies. *Demography*, 54(2), 603–630.
- De Bonis, V., & Spataro, L. (2005). Taxing capital income as pigouvian correction: The role of discounting the future. *Macroeconomic Dynamics*, 9, 469–477.
- De Bonis, V., & Spataro, L. (2010). Social discounting, migration and optimal taxation of savings. *Oxford Economic Papers*, 62, 603–623.
- de la Croix, D., & Michel, P. (2002). *A theory of economic growth*. Cambridge: Cambridge University Press.
- Erosa, A., & Gervais, M. (2002). Optimal taxation in life-cycle economies. *Journal of Economic Theory*, 105, 338–369.
- Farhi, E., & Werning, I. (2005). Inequality, social discounting and estate taxation. NBER Working Paper 1148.
- Garriga, C. (2003). *Optimal fiscal policy in overlapping generations models*. New York: Mimeo.

- Golosov, M., Kocherlakota, N., & Tsyvinski, A. (2003). Optimal indirect and capital taxation. *Review of Economic Studies*, 70, 569–587.
- Grossman, H. I., & Van Huyck, J. B. (1988). Sovereign debt as a contingent claim: Excusable default, repudiation, and reputation. *American Economic Review*, 78(5), 1088.
- Halvorsen, E., & Thoresen, T. O. (2017). Distributional effects of the wealth tax under a lifetime-dynastic income concept. CESifo Working Paper 6614.
- Judd, K. L. (1985). Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics*, 28, 59–83.
- Judd, K. L. (1999). Optimal taxation and spending in general competitive growth models. *Journal of Public Economics*, 71, 1–26.
- Kanbur, R., & Stiglitz, J. E. (2015). Dynastic inequality, mobility and equality of opportunity. Centre for Economic Policy Research Discussion Paper 10542.
- Kocherlakota, N. (2005). Zero expected wealth taxes: A mirrlees approach to dynamic optimal taxation. *Econometrica*, 73, 1587–1621.
- Lansing, K. (1999). Optimal redistributive capital taxation in a neoclassical growth model. *Journal of Public Economics*, 73(3), 423–453.
- Michel, P. (1990). Criticism of the social time-preference hypothesis in optimal growth. Catholique de Louvain—Center for Operations Research and Economics Paper 9039.
- Occhino, F. (2008). Optimal fiscal policy when migration is feasible. *B.E. Journal of Economic Analysis & Policy*, 8, 1682–1935.
- Piketty, T. (2014). *Capital in the twenty-first century*. Cambridge: Harvard University Press.
- Ramsey, F. P. (1927). A contribution to the theory of taxation. *Economic Journal*, 37, 47–61.
- Ramsey, F. P. (1928). A mathematical theory of saving. *Economic Journal*, 38, 543–559.
- Razin, A., Sadka, E., & Swagel, P. (2002). Tax burden and migration: A political economy theory and evidence. *Journal of Public Economics*, 85, 167–190.
- Reinhorn, L. J. (2013). *On optimal redistributive capital taxation*. New York: mimeo.
- Reis, C. (2012). Social discounting and incentive compatible fiscal policy. *Journal of Economic Theory*, 147, 2469–2482.
- Solé-Auró, A., & Crimmins, E. M. (2008). Health of immigrants in European countries. *International Migration Review*, 42(4), 861–876.
- Spataro, L., & De Bonis, V. (2008). Accounting for the “disconnectedness” of the economy in OLG models: A case for taxing capital income. *Economic Modelling*, 25, 411–421.
- Stiglitz, J. E. (2015). In praise of Frank Ramsey’s contribution to the theory of taxation. *Economic Journal*, 125(583), 235–268.
- Straub, L., & Werning, I. (2014). *Positive long run capital taxation: Chamley–Judd revisited*. (No. w20441). National Bureau of Economic Research.
- Weil, P. (1989). Overlapping families of infinitely lived agents. *Journal of Public Economics*, 38, 183–198.
- Zhu, X. (1992). Optimal fiscal policy in a stochastic growth model. *Journal of Economic Theory*, 58, 250–289.