Efficient Subsidization of Human Capital Accumulation with Overlapping Generations and Endogenous Growth

by

Wolfram F. Richter and Christoph Braun

TU Dortmund University

April 2009

First Draft to be presented at the
Conference in Honour of Robin Boadway
Queen's University, Kingston, Ontario, 14-15 May 2009

Abstract:

This paper studies second-best efficient policies for human capital in an OLG model with endogenous growth. When young, individuals decide on education, saving and nonqualified labour. When old individuals supply qualified labour. Growth equilibria are inefficient in *laissez-faire* because government expenditures have to be financed by distortionary taxes. The inefficiency is exacerbated if selfish individuals externalize the positive effect of education on descendents' productivity. It is shown to be second best to subsidize human capital accumulation even relative to the first best. This holds for selfish individuals at balanced growth and it equally holds for altruists in the dynasty's first generation.

Keywords: OLG model, endogenous growth, endogenous labour, education and saving, intergenerational externalities, optimal taxation

Address:

Wolfram F. Richter and Christoph Braun, TU Dortmund University, Department of Economics, 44221 Dortmund, Germany. Corresponding author: Wolfram.Richter@tu-dortmund.de, phone: +49-231-755-3146, fax: +49-231-755-5404

1. Introduction

The implications of adding endogenous education to the Ramsey problem of efficient taxation are by now well understood. It is however irritating to learn that they strongly depend on whether the representative taxpayer lives for finite or infinite periods. If the taxpayer's planning horizon is infinite, the reason for employing distortionary linear taxes turns out to be weak. This point has been made first by Chamley (1986) and Judd (1985) and it extends to the model with endogenous education as has been demonstrated by Bull (1993), Jones, Manuelli, Rossi (1993, 1997), and Atkeson, Chari, Kehoe (1999). Along a balanced growth path no use should be made of distortionary taxes. The problem one may however have with this kind of result is that it is of little help in identifying differences between human and nonhuman capital to be taken into account by tax policy.

The existence of such differences is suggested when solving the finite Ramsey tax problem. It then becomes clear that the efficiency of not taxing saving is primarily a reflection of the taxpayer's intertemporal preference structure. In particular, savings should not be taxed if the taxpayer's utility is weakly separable between consumption and labour and homothetic in consumption (Atkinson and Stiglitz, 1972; Sandmo, 1974). By way of contrast, the design of efficient education policy is more a reflection of the specific properties of the earnings function. If the earnings function is weakly separable in qualified labour supply and education and if the elasticity with respect to the latter is constant, then the choice of education remains undistorted in the second-best optimum. In other words, the return to education should equal its cost before taxes and subsidies (Jacobs and Bovenberg, 2008; Bovenberg and Jacobs, 2005; Richter, 2006). Furthermore, labour should be taxed such that qualified labour is distorted less than nonqualified labour (Richter, 2008).

Studying optimal tax design in the finite Ramsey problem is known to suffer from various shortcomings. The most obvious ones are the focussing on a representative taxpayer and the ignoring of heterogeneity and informational asymmetry. Critical is also the ignoring of potential reasons of capital market or policy failure. The present paper however ignores all such shortcomings. Its sole objective is to contribute to the attempt of overcoming the static nature of the finite Ramsey model. More specifically, the paper explores the implications for second-best efficient policies when acknowledging the fact that human capital accumulation is a dynamic activity affecting the well-being of descendent generations. If descendent generations benefit by the human capital investments of preceding generations, one would not necessarily expect non-distortionary education policy to be efficient. The differentiation of qualified and nonqualified labour shaping efficient taxation in the finite Ramsey problem raises additional questions when applied to a scenario in which qualified parents and nonqualified children coexist. Hence it is not clear to what extent the results of the static Ramsey analysis survive in a dynamic framework with overlapping generations.

The present paper studies second-best efficient policies for education, labour and saving in a two-period overlapping-generations model with endogenous growth. Individuals are assumed to supply nonqualified labour when young and qualified labour when old. They may be either perfect altruists with respect to descendent generations or they may behave selfishly. The implications of selfishness have been studied before by Docquier, Paddison and Pestieau (2006) for a framework in which the government is not constrained in the use of policy instruments. The authors show that decentralizing the first best requires subsidizing education. The present study goes beyond Docquier et al. (2006) by endogenizing labour supply and by assuming that the government can only employ linear taxes and subsidies on labour and education. As it turns out it is second best along balanced growth to encourage education to such an extent that human capital accumulation is positively distorted relative to the first best. This means that the marginal social cost of human capital should exceed the marginal social return in the long-run second-best optimum. This is a striking result. Striking is the efficiency of distortion as such and even more striking is the sign of the distortion. Not surprising is the need to subsidize education relative to the *laissez-faire*. This is so as the intergenerational externalities of human capital investments have to be internalized. *A priori* it is not obvious however why the accumulation of human capital should be distorted relative to the *first best*. The sign of the distortion may wonder even more. Subsidizing education requires government revenue which in the model has to be raised by distortionary taxes on labour and savings. With the intuition of Lipsey and Lancaster (1956/57) in mind one might hypothesize that it is second best to provide insufficient incentives for human capital investment if labour has to be taxed and if the level of comparison is the first best. The contrary is however true. The reason is that taxes on labour have a negative effect on education and growth and that human capital policy has to compensate for this dynamic effect. More precisely, it will be demonstrated that the strength by which human capital accumulation should be positively distorted in the long run increases in three factors. One factor is the social marginal utility of income, the second factor will be called the dynamic cost of education and the third factor is the gap between the marginal return to capital and the rate of balanced growth. In other words, the more binding the non-availability of lump-sum taxes is and the more costly human capital accumulation is and the more deficient growth is, the more should human capital accumulation be encouraged beyond what appears to be first best. Additionally it is shown to be second best to distort qualified labour less than nonqualified labour. Hence the results derived for the two-period Ramsey tax model carry over to the overlapping-generations model only in part. The structural design of efficient labour taxation is preserved while the education policy differs strongly.

Assuming altruistic individuals changes some conclusions but not all. Altruists internalize the positive effect that education has on descendents' productivity. Hence the need for government intervention is reduced. However, the second source of inefficiency modelled in this paper does not vanish. The second source is the need to employ distortionary taxes for financing exogenous government expenditures. The implications for second-best policy are shown to differ markedly between the first generation and all descendent generations. With respect to descendent generations the following results are obtained. Qualified and nonqualified labour should be taxed uniformly and the accumulation of human capital should not be distorted. Such results strongly contrast with those derived for the case of selfish individuals. Against these, the results obtained for the first generation are more similar. The specific properties of optimal policy for the first generation depend on initial values. After neutralizing the impact of initialization a case can however be made for subsidizing the human capital investment. The reason is the same encountered when individuals are assumed to be selfish. Taxing labour has a negative effect on education and growth and human capital policy has to compensate for this dynamic effect. The unifying bottom line for selfish and altruistic individuals is as follows. Altruism well reduces the need to subsidize education relative to *laissez-faire* and altruism also implies that descendent generations should have non-distorted incentives to accumulate human capital. The short-run policy recommendations for altruism however agree with the long-run recommendations for selfishness. Labour has to be taxed and the resulting decrease in growth has to be compensated by subsidizing human capital accumulation relative to the first-best. Whether saving should be taxed or not, primarily depends on assumptions made with regard to consumption preferences.

The paper is structured as follows. Section 2 sets up the model of a two-period overlapping-generations model with endogenous growth. The first-order conditions characterizing solutions of the planner's first-best maximization are derived. Section 3 studies the planner's problem when individuals behave selfishly and when only linear policy instruments are available. Section 4 studies the same problem for individuals which are altruistic towards descendent generations. Section 5 summarizes.

2. The model and the planner's first-best problem

Consider a representative individual of generation *t*. The individual is living for two periods. Lifetime utility is given by $U(C_{0t}, C_{1t} L_{0t}, L_{1t})$ where the arguments $C_{0t}, C_{1t} L_{0t}, L_{1t}$ denote consumption and non-leisure in life-cycle periods zero and one. Non-leisure time L_{1t} is identical with second-period labour supply. By contrast, only $L_{0t} - E_t$ is time spent in the market while time E_t is spent on education. The effect of education is to increase human capital and labour productivity. H_{t-1} is the stock of human capital inherited by generation *t* from the preceding one. H_t is the stock built up by generation *t* and effective in *t*'s second life period. Human capital accumulates according to the equation

$$[G(E_t) + 1 - \delta_H] H_{t-1} = H_t . \qquad (\mu_t \beta^t) \qquad (1)$$

Hence it depreciates at the rate of δ_H and it increases by the endogenous choice of E_t. The learning function $G(E_t)$ is assumed to display positive but diminishing returns, *G'>0>G"* with *G*(0)=0. The elasticity of education, $\eta \equiv EG'/G$, is then positive but smaller than one. $\mu_t \beta^t$ is a Lagrange multiplier associated with the planner's problem we are about to set up. There is a second stock variable K_t to be interpreted as (nonhuman) capital built up by generation *t* in their first life period. It depreciates at the rate of δ_K. Production displays constant returns to capital and effective labour. The *resource constraint* is

$$F(K_{t-1}, (L_{0t} - E_t) H_{t-1}, L_{1t-1} H_{t-1}) + (1 - \delta_K) K_{t-1}$$

$$= C_{0t} + C_{1t-1} + f_t E_t H_{t-1} + K_t + A_t \qquad (\alpha_t \beta^t) \qquad (2)$$

where A_t denotes exogenous government expenditure. Obviously, human capital accumulation is labour augmenting. When taking partial derivatives use is made of the following short-forms:

$$F_{Kt-1} \equiv \frac{\partial F}{\partial K_{t-1}}, \; F_{0t} \equiv \frac{\partial F}{\partial ((L_{0t} - E_t) H_{t-1})}, \; F_{1t-1} \equiv \frac{\partial F}{\partial (L_{1t-1} H_{t-1})}$$

It is suggestive to interpret $L_{0t} - E_t$ as nonqualified labour and L_{1t} as qualified labour. The two kinds of labour may be perfect or imperfect substitutes in production. The return to education is increased productivity. Two kinds of cost are modelled. There is the cost of foregone earnings captured by $F_{0t}\ E_t H_{t-1}$ and there are direct costs which for simplicity sake are modelled as a linear function of learning time and inherited human capital, $f_t E_t H_{t-1}$.

The planner maximizes

$$\sum_{t=0}^{\infty} \beta^t U(C_{0t}, C_{1t} L_{0t}, L_{1t}) \tag{3}$$

in $C_{0t}, C_{1t} L_{0t}, L_{1t}, E_t, H_t, K_t$ (t=0,1, ..) subject to (1) and (2). The parameters $K_{-1}, H_{-1}, L_{1-1} = L_{1,t=-1}$ are exogenously given. The first-order conditions are as follows:

$$U_{C0t} = \alpha_t\,,\quad U_{C1t} = \alpha_{t+1}\beta\,,\quad F_{0t}\ H_{t-1}\, U_{C0t} = -U_{L0t}\,,\quad F_{1t} H_t\, U_{C1t} = -U_{L1t}\,, \tag{4}$$

$$F_{Kt} + 1 - \delta_K\ = U_{C0t} / U_{C1t}\ = U_{C0t} / \beta\, U_{C0t+1}\,, \tag{5}$$

$$\mu_t G_t^{'} = \alpha_t\ (f_t + F_{0t})\,, \tag{6}$$

$$\alpha_{t+1}\beta\,[\,F_{1t} L_{1t} + F_{0t+1} \cdot (\,L_{0t+1} - E_{t+1}) - f_{t+1} E_{t+1}\,] = \mu_t - \beta\ [G_{t+1} + 1 - \delta_H\,]\ \mu_{t+1} \tag{7}$$

Conditions (4) characterize efficient consumption and labour choices. Condition (5) characterizes efficient saving and efficient capital. Condition (6) characterizes the efficient choice of E_t and (7) is the condition characterizing the efficient choice of H_t. Solving (6) for μ_t and inserting into (7) yields after some straightforward manipulations the condition characterizing the *efficient accumulation of human capital*,

$$F_{1t} L_{1t} + F_{0t+1}\ L_{0t+1} - (\,F_{0t+1} + f_{t+1}\,)\,E_{t+1}$$

$$=\ [\,F_{Kt} + 1 - \delta_K\,]\,\frac{f_t + F_{0t}}{G_t^{'}}\ -\ [G_{t+1} + 1 - \delta_H\,]\ \frac{f_{t+1} + F_{0t+1}}{G_{t+1}^{'}}\,. \tag{8}$$

The first term on the left-hand side, $F_{1t} L_{1t}$, is the return to human capital accruing to generation *t* in the second period of life and the difference $F_{0t+1}\ L_{0t+1} - (\,F_{0t+1} + f_{t+1}\,)\,E_{t+1}$ is the return accruing to individuals of the next generation in their first life period. $\frac{f_t + F_{0t}}{G_t^{'}}$ is the cost of human capital in period *t* and $\frac{f_{t+1} + F_{0t+1}}{G_{t+1}^{'}}$ is the cost of human capital one period later. The right-hand side of (8) is the cost resulting from investing in period *t* instead of postponing the investment to the next period. When separating terms indexed by *t* from terms indexed by *t*+1 the efficiency condition (8) implies

$$F_{1t}L_{1t} - [F_{Kt} + 1 - \delta_K]\frac{f_t + F_{0t}}{G_t^{'}} = -F_{0t+1}\,L_{0t+1} - (F_{0t+1} + f_{t+1})E_{t+1}[\frac{G_{t+1} + 1 - \delta_H}{E_{t+1}G_{t+1}^{'}} - 1]$$

$$< -F_{0t+1}\,L_{0t+1} - (F_{0t+1} + f_{t+1})E_{t+1}[\frac{1}{\eta_{t+1}} - 1] < 0.$$

The first inequality follows from $\frac{G_{t+1} + 1 - \delta_H}{E_{t+1}G_{t+1}^{'}} > \frac{G_{t+1}}{E_{t+1}G_{t+1}^{'}}$ while the second inequality relies on the regressivity of the learning function, $\eta_{t+1} < 1$. Hence it is first-best that generation *t*'s cost, $[F_{Kt} + 1 - \delta_K]\frac{f_t + F_{0t}}{G_t^{'}}$, exceeds generation *t*'s return to human capital, $F_{1t}L_{1t}$. The difference is the positive external effect on generation *t*+1 which has to be internalized by first-best policy when individuals are selfish.

Along a *balanced growth path* L_{0t}, L_{1t} and E_t are constant while consumption, output and both types of capital all grow at the common rate of $G + 1 - \delta_H \equiv g$ which, by assumption, is strictly lower than the gross rate of return to capital, $F_K + 1 - \delta_K$ ("*condition of transversality*"). The marginal productivities F_K, F_i (*i*=0,1) are homogeneous of degree zero. A balanced growth path requires marginal utilities of consumption U_{Cit} being constant in *t*=0,1,… for each *i*=0,1. This is guaranteed whenever the utility function is linear homogeneous in consumption. In this case, (5) implies

$$1/\beta = F_K + 1 - \delta_K > G + 1 - \delta_H \equiv g\,. \tag{9}$$

Evaluated at a balanced growth path, (8) simplifies to

$$F_1L_1 + F_0\,L_0 - (F_0 + f)E = [(F_K - \delta_K) - (G - \delta_H)]\frac{f + F_0}{G'}\,. \tag{10}$$

The following analysis studies the question of whether it is second best to provide or not to provide efficient incentives for human capital accumulation. As we shall see, much depends on individual behaviour and the question of whether individuals are perfect altruists towards their children or not. In the altruistic model – also called the dynasty model – individuals are assumed to maximize (3). In the other case which is considered first the representative individual is assumed to maximize own lifetime utility

$$U(C_{0t}, C_{1t} L_{0t}, L_{1t}) \tag{11}$$

subject to the own lifetime budget constraint. We study both scenarios and we start by analyzing efficient taxation in the standard OLG framework with selfish individuals.

3. Optimal taxation in the standard OLG model with selfish individuals

The selfish individual representing generation *t* is assumed to maximize (11) in the six choice variables $C_{0t}, C_{1t} L_{0t}, L_{1t}, E_t, K_t$ subject to the lifetime budget constraint

$$\pi_t \omega_{0t}(L_{0t} - E_t) H_{t-1} + \pi_{t+1}\omega_{1t} L_{1t} H_t$$

$$= \pi_t C_{0t} + \pi_{t+1} C_{1t} + \pi_t \varphi_t E_t H_{t-1} + (\pi_t - R_{t+1}\pi_{t+1}) K_t \qquad (\lambda_t) \tag{12}$$

$$H_t = H_t(E_t) \equiv [G(E_t) + 1 - \delta_H] H_{t-1} \; . \tag{13}$$

In this optimization H_{t-1} is treated as an exogenous parameter. For each *t* there are six first-order conditions

$$U_{C0t} = \pi_t \lambda_t, \quad U_{C1t} = \pi_{t+1}\lambda_t, \quad \omega_{0t} \; H_{t-1} \, U_{C0t} = -U_{L0t}, \quad \omega_{1t} H_t \, U_{C1t} = -U_{L1t}, \tag{14}$$

$$\omega_{1t} L_{1t} \, G_t^{'} \, U_{C1t} = (\varphi_t + \omega_{0t}) \, U_{C0t}, \quad R_{t+1} = \pi_t / \pi_{t+1} \tag{15}$$

which can be used to substitute for the five relative prices $\omega_{0t}, \omega_{1t}, \varphi_t, \pi_{t+1}/\pi_t, R_{t+1}$ and the Lagrange multiplier λ_t. After substituting, the budget constraint (12) can be written as

$$\sum_{i=0}^{1} [C_{it} U_{Cit} + L_{it} U_{Lit}] = L_{1t} U_{L1t} \; G_t^{'} \; E_t \; H_{t-1} / H_t \qquad (\tilde{\lambda}_t \beta^t) \tag{16}$$

which will take the role of an *Incentive* (or *Implementation*) *Constraint* in the planner's second-best problem. Because of

$$(\varphi_t + \omega_{0t}) \; E_t H_{t-1} = -\frac{L_{1t} U_{L1t}}{U_{C0t}} \frac{G_t^{'} E_t}{H_t} \; H_{t-1}, \tag{17}$$

the right-hand side of (16) can be interpreted as the private value of the cost of education. This cost depends on various factors. As it turns out, the dependence on E_t - measured by the marginal variation in H_t - is of particular significance when characterizing second-best policy. Let us call the resulting marginal variation the *private marginal cost of human capital accumulation* or, for short, the *dynamic cost of education*. The formal definition is

$$MC_t^{HC} \equiv -\frac{d}{dH_t}\left[\frac{L_{1t}U_{L1t}}{U_{C0t}}\frac{G_t^{'}E_t}{H_t}H_{t-1}\right] = -\frac{L_{1t}U_{L1t}}{U_{C0t}}\left(\frac{\partial}{\partial H_t}+\frac{dE_t}{dH_t}\frac{\partial}{\partial E_t}\right)\left[\frac{G_t^{'}E_t}{H_t}\right]H_{t-1}$$

$$= -\frac{L_{1t}U_{L1t}}{U_{C0t}}\frac{1}{H_t}\left[-\frac{G_t^{'}E_t}{H_t}H_{t-1}+1+\frac{G_t^{''}E_t}{G_t^{'}}\right]. \qquad (18)$$

In order to guarantee that MC_t^{HC} is positive, the bracketed expression on the right-hand side of (18) must be positive. Such positivity is ensured whenever $G(E_t)$ is isoelastic, i.e. if $0<\eta = E_t G_t^{'}/G_t$ = constant in E_t. In this special case, $-\frac{G_t^{'}E_t}{H_t}H_{t-1}+1+\frac{G_t^{''}E_t}{G_t^{'}} = \eta\frac{1-\delta_H}{G_t+1-\delta_H}$ > 0. In what follows, MC_t^{HC} is assumed to be positive.

The planner maximizes (3) in $C_{0t}, C_{1t}, L_{0t}, L_{1t}, E_t, H_t, K_t$ (t=0,1, ..) subject to (16), (1), and (2). The solutions are second best in the sense that they have to fulfil the Incentive Constraint (16) in addition to the first-best constraints (1) and (2). If lump-sum taxes were available, the planner could ignore (16). Inclusion of (16) in the set of constraints implies that the planner is restricted in the choice of policy instruments. The restriction is however not an arbitrary one. Quite to the contrary, implicit in the derivation of (16) is the assumption that the planner is not constrained in setting consumer prices $\omega_{0t}, \omega_{1t}, \varphi_t, R_{t+1}$. This means in particular that labour income can be taxed at different rates over an individual's lifecycle. If such differentiation is ruled out by assumption, the planner has to respect an additional constraint which may have strong implications for the design of optimal taxation. See Erosa and Gervais (2002) for a discussion of this point in an OLG Model without endogenous education.

To solve the planner's problem set

$$W_t \equiv U_t + \tilde{\lambda}_t\left\{\sum_{i=0}^{1}[C_{it}U_{Cit}+L_{it}U_{Lit}] - L_{1t}U_{L1t}\,G_t^{'}\,E_t\,H_{t-1}/H_t\right\}. \qquad (19)$$

When discussing the efficient taxation of saving, particular focus is on utility functions which are weakly separable between consumption and non-leisure,

$$U = U(V(C_0, C_1), L_0, L_1), \qquad (20)$$

with a linear homogeneous nested function *V*. Such utility functions are known to have the attractive property that the private and the social marginal rates of substitution in consumption are equal.

Remark: Assume weak separability between consumption and non-leisure and assume linear homogeneity of *V*. Then

$$\frac{W_{C0t}}{W_{C1t}} = \frac{U_{C0t}}{U_{C1t}}. \tag{21}$$

The proof is straightforward. For the sake of simplicity the index *t* is suppressed.

$$\frac{W_{Ci}}{U_{Ci}} = \frac{1}{U_{Ci}}\frac{d}{dC_i}\{U+\tilde{\lambda}\sum_{j=0}^{1}[C_jU_{Cj}+L_jU_{Lj}] - \tilde{\lambda}\, L_1U_{L1}\, G^{'}E\,H_{-1}/H_0\}$$

$$= 1+\tilde{\lambda}\,\{1+\sum_{j=0}^{1}[C_j\frac{U_{CjCi}}{U_{Ci}}+L_j\frac{U_{LjCi}}{U_{Ci}}] - L_1\frac{U_{L1Ci}}{U_{Ci}}\,G^{'}E\,H_{-1}/H_0\}$$

$$= 1+\tilde{\lambda}\,\{1+V\frac{U_{VV}}{U_V}+\sum_{j=0}^{1}L_j\frac{U_{VLj}}{U_V} - L_1\frac{U_{VL1}}{U_V}\,G^{'}E\,H_{-1}/H_0\} = \text{constant in } i\text{=0,1.}$$

The first question addressed is the one raised by efficient taxation of saving. The relevant first-order conditions associated with the planner's problem are as follows:

$$\frac{\partial}{\partial C_{0t}},\ \frac{\partial}{\partial L_{0t}}:\quad W_{C0t}=\alpha_t=-\frac{W_{L0t}}{F_{0t}H_{t-1}} \tag{22}$$

$$\frac{\partial}{\partial C_{1t}},\ \frac{\partial}{\partial L_{1t}}:\quad W_{C1t}=\alpha_{t+1}\,\beta=-\frac{W_{L1t}}{F_{1t}H_t} \tag{23}$$

$$\frac{\partial}{\partial K_{t+1}}:\ \alpha_{t+1}\beta\,[\,F_{Kt}+1-\delta_K\,] = \alpha_t \tag{24}$$

By (21),

$$F_{Kt}+1-\delta_K = \frac{\alpha_t}{\alpha_{t+1}\beta} = \frac{W_{C0t}}{W_{C1t}} = \frac{U_{C0t}}{U_{C1t}}\ . \tag{22}$$

Proposition 1: Assuming selfish behaviour, weak separability between consumption and non-leisure and assuming linear homogeneity of *V*, then it is second-best efficient not to distort saving.

A slight variant of Proposition 1 is obtained when focussing on balanced growth. Hence assume $L_{it}=L_i$ (*i*=0,1) and E_t=*E* to be constant across time and H_t, K_t, C_{it} (*i*=0,1) to grow at the common gross rate $G+1-\delta_H=g$, so that $H_{t-1}=g^t H, K_{t-1}=g^t K, C_{it}=g^t C_i$ results. Assume furthermore that utility *U* is linear homogeneous in consumption. *W* as defined by (19) is then equally linear homogeneous in consumption just as U_{Li} and W_{Li} while U_{Ci}, W_{Ci} are functions which are homogeneous of degree zero in consumption. As a result, the growth factor g^t cancels out of condition (22). After cancelling out, the only variable carrying an index *t* in $W_{C0t}=-\dfrac{W_{L0t}}{F_{0t}H_{t-1}}$ is $\tilde{\lambda}_t$ which therefore must be constant along the balanced growth path, $\tilde{\lambda}_t=\tilde{\lambda}>0$. The same holds for $\alpha_t=\alpha$, $U_{C0t}=U_{C0}$. The private rate of return to saving,

$$R_{t+1}=\frac{\pi_t}{\pi_{t+1}}=\frac{1}{\beta}\frac{U_{C0t}}{U_{C0t+1}}=\frac{1}{\beta}$$

then equals the social rate of return to capital, $F_K+1-\delta_K$.

Proposition 2: Assume selfish behaviour and *U* to be linear homogeneous in consumption. At a balanced growth path it is second-best efficient not to distort saving.

Propositions 1 and 2 extend earlier results of Atkinson and Stiglitz (1972), Sandmo (1974), Atkeson, Chari and Kehoe (1999) and others to the present framework.

Turn next to human capital and education. We are going to prove that it is efficient to encourage human capital investment along a balanced growth path to such an extent that the first-best level is exceeded. To show this, evaluate the first-order conditions with respect to E_t and H_t:

$$\frac{\partial}{\partial E_t}: \quad \mu_t\, G_t^{'}\, H_{t-1} = \tilde{\lambda}_t \frac{L_{1t}U_{L1t}}{H_t}[G_t^{'}+G_t^{''}E_t]H_{t-1} + \alpha_t(f_t+F_{0t})H_{t-1} \Rightarrow$$

$$\frac{\mu_t}{\alpha_t} = \frac{f_t + F_{0t}}{G_t^{'}} + \frac{\tilde{\lambda}_t}{\alpha_t}\,\frac{L_{1t}U_{L1t}}{H_t}[1+\frac{G_t^{''}E_t}{G_t^{'}}] \tag{25}$$

$$\frac{\partial}{\partial H_t}:\ \tilde{\lambda}_t\, L_{1t}U_{L1t}\, G_t^{'}\, E_t\, H_{t-1}/H_t^2 - \tilde{\lambda}_{t+1}\,\beta\, L_{1t+1}U_{L1t+1}\, G_{t+1}^{'}\, E_{t+1}/H_{t+1}$$

$$+\ \alpha_{t+1}\,\beta\,[F_{1t}L_{1t} + F_{0t+1}\cdot(L_{0t+1} - E_{t+1}) - f_{t+1}E_{t+1}] + \mu_{t+1}\,\beta\,[G_{t+1} + 1 - \delta_H] = \mu_t \tag{26}$$

Making use of (24) and (25) we obtain

$$[F_{1t}L_{1t} + F_{0t+1}\cdot(L_{0t+1} - E_{t+1}) - f_{t+1}E_{t+1}] + \frac{f_{t+1} + F_{0t+1}}{G_{t+1}^{'}}(G_{t+1} + 1 - \delta_H)$$

$$+\ \frac{\tilde{\lambda}_{t+1}}{\alpha_{t+1}}\,\frac{L_{1t+1}U_{L1t+1}}{H_{t+1}}[1+\frac{G_{t+1}^{''}E_{t+1}}{G_{t+1}^{'}} - \frac{G_{t+1}^{'}E_{t+1}}{H_{t+1}}\,H_t](G_{t+1} + 1 - \delta_H)$$

$$= [\frac{f_t + F_{0t}}{G_t^{'}} + \frac{\tilde{\lambda}_t}{\alpha_t}\,\frac{L_{1t}U_{L1t}}{H_t}(1+\frac{G_t^{''}E_t}{G_t^{'}} - \frac{G_t^{'}E_t}{H_t}\,H_{t-1})][F_{Kt} + 1 - \delta_K]. \tag{27}$$

Making use of (18), condition (27) can be written in a more structured form:

$$[F_{1t}L_{1t} + F_{0t+1}\cdot(L_{0t+1} - E_{t+1}) - f_{t+1}E_{t+1}]$$

$$+\ [\frac{f_{t+1} + F_{0t+1}}{G_{t+1}^{'}} - \frac{\tilde{\lambda}_{t+1}}{\alpha_{t+1}}\,U_{C0t+1}\cdot MC_{t+1}^{HC}](G_{t+1} + 1 - \delta_H)$$

$$= [\frac{f_t + F_{0t}}{G_t^{'}} - \frac{\tilde{\lambda}_t}{\alpha_t}\,U_{C0t}\cdot MC_t^{HC}][F_{Kt} + 1 - \delta_K]. \tag{28}$$

Setting

$$\Delta_t \equiv \frac{\tilde{\lambda}_t}{\alpha_t}\,U_{C0t}\cdot MC_t^{HC}\cdot(F_{Kt} + 1 - \delta_K) - \frac{\tilde{\lambda}_{t+1}}{\alpha_{t+1}}\,U_{C0t+1}\cdot MC_{t+1}^{HC}\cdot(G_{t+1} + 1 - \delta_H) \tag{29}$$

allows one to restate (28):

$$\Delta_t = \frac{f_t + F_{0t}}{G_t^{'}}(F_{Kt} + 1 - \delta_K) - \frac{f_{t+1} + F_{0t+1}}{G_{t+1}^{'}}(G_{t+1} + 1 - \delta_H)$$

$$- F_{1t}L_{1t} - [F_{0t+1}\cdot(L_{0t+1} - E_{t+1}) - f_{t+1}E_{t+1}].$$

Comparing this condition with (8) shows that Δ_t is the wedge by which human capital accumulation should be efficiently distorted. A positive wedge stands for subsidization relative to the first best. Condition (29) invites to be evaluated at a balanced growth path. By the reasons given in the proof of Proposition 2, $\tilde{\lambda}_t = \tilde{\lambda} > 0$, $\alpha_t = \alpha$, $U_{C0t} = U_{C0}$, $MC_t^{HC} = MC^{HC}$ has to hold. Setting $R = F_K + 1 - \delta_K$ and evaluating (29) at a balanced growth path we end up with

$$\Delta = \frac{\tilde{\lambda}}{\alpha} U_{C0} \cdot MC^{HC} \cdot (R - g). \tag{30}$$

Interpret $\frac{\tilde{\lambda}}{\alpha} U_{C0} > 0$ as the social marginal utility of income measured in real terms. This factor is positive if the Incentive Constraint is binding, $\tilde{\lambda} > 0$, which is the case if the non-availability of lump-sum taxes is a binding constraint. In this sense the factor measures the cost resulting from the non-availability of lump-sum taxes. MC^{HC} is the dynamic cost of education which is positive by assumption. Finally, $R - g$ is the growth gap. By the transversality condition (9) the gap must be positive as well. Hence Δ is the product of three factors each of which is positive.

Proposition 3: Assume selfish behaviour and *U* to be linear homogeneous in consumption. At a balanced growth path it is second-best efficient to subsidize human capital accumulation relative to the first best. The strength of positive distortion increases in (i) the cost resulting from the non-availability of lump-sum taxes, (ii) the dynamic cost of education, and (iii) the growth gap.

This is a striking result for reasons explained before. *A priori* it is not clear that human capital accumulation should be distorted along balanced growth while capital accumulation should not be distorted subject to appropriately chosen utility functions. The sign of the efficient distortion is neither obvious when taking the first best as the standard of comparison. It is rather evident and has been noted before that the *laissez-faire* level of education is inefficient from the first-best perspective. Without government intervention selfish individuals externalize the positive effect of own education on descendent generations' welfare. The critical question therefore is in what direction second-best policy should deviate from the first best if such a deviation is efficient. Note that any revenue needed to subsidize the monetary cost of education has to be raised by distortionary labour taxes. With the intuition of Lipsey and Lancaster (1956/57) in mind one could have hypothesized that it is second best to give negative incentives for human capital accumulation relative to the first best if labour has to be taxed. The contrary is however true. The intuition is as follows. The tax on labour has a depressing effect on education. We speak of a depressing effect because it does not necessarily mean a statically inefficient choice of education. In static analysis it may well be second-best efficient to distort only labour but not education (Jacobs and Bovenberg, 2008; Bovenberg and Jacobs, 2005; Richter, 2006). Still, the choice of education will be reduced when labour income is taxed. In the static second-best optimum, education is reduced in the same proportion as nonqualified labour and even more than qualified labour (Richter, 2008). In dynamic analysis this reduction slows down long-run growth. A dynamic inefficiency results the strength of which increases in the factors listed by Proposition 2. Education policy has to compensate for this effect.

What can be said about the efficient taxation of nonqualified labour relative to qualified labour? We only analyze this question for balanced growth. Setting $\eta = EG'/G$ without assuming constancy of η, (19) can be written as

$$W \equiv U + \tilde{\lambda}\,\{\sum_{i=0}^{1}[C_i U_{Ci} + L_i U_{Li}] - L_1 U_{L1}\,\eta\, G/g\}\ . \tag{31}$$

Note that the specification of *W* according to (31) is structurally asymmetric in L_0 and L_1. This indicates that qualified and nonqualified labour should be taxed differently. To be more specific evaluate (22) and (23). The former condition states $W_{L0} + F_0 H\, W_{C0} = 0.$ $\Leftrightarrow$

$$(1+\tilde{\lambda})[U_{L0} + F_0 H\, U_{C0}]$$

$$= \tilde{\lambda}\,\{\frac{\eta G}{g}\, L_1\,(U_{L1L0} + F_0 H\, U_{L1C0}) - \sum_{i=0}^{1}[C_i U_{CiL0} + L_i U_{LiL0}]$$

$$- F_0 H \sum_{i=0}^{1}[C_i U_{CiC0} + L_i U_{LiC0}]\}$$

$$= -\tilde{\lambda}\,\{\sum_{i=0}^{1} C_i U_{CiL0} + L_0 U_{L0L0} + (1 - \frac{\eta G}{g})\, L_1 U_{L1L0} + b_0\, F_0 H\, U_{C0}\}$$

with $b_0 \equiv [\sum_{i=0}^{1} C_i U_{CiC0} + L_0 U_{L0C0} + (1 - \frac{\eta G}{g})\, L_1 U_{L1C0}]/U_{C0}$

$\Leftrightarrow$

$$U_{L0} + F_0 H\, U_{C0} = -\,a_0\,[\sum_{i=0}^{1} C_i U_{CiL0} + L_0 U_{L0L0} + (1 - \frac{\eta G}{g})\, L_1 U_{L1L0} - b_0\, U_{L0}] \tag{32}$$

with $a_0 \equiv \tilde{\lambda}/[1+\tilde{\lambda}+b_0\,\tilde{\lambda}]$.

Similarly, (23) $\Leftrightarrow$ $W_{L1}+F_1 gH\, W_{C1}=0$ $\Leftrightarrow$

$(1+\tilde{\lambda})[U_{L1}+F_1 gH\, U_{C1}]$

$$= \tilde{\lambda}\{\frac{\eta G}{g}[U_{L1}+L_1(U_{L1L1}+F_1 gH\, U_{L1C1})]-\sum_{i=0}^{1}[C_i U_{CiL1}+L_i U_{LiL1}]$$

$$-F_1 gH\sum_{i=0}^{1}[C_i U_{CiC1}+L_i U_{LiC1}]\}$$

$$= -\tilde{\lambda}\{\sum_{i=0}^{1}C_i U_{CiL1}+L_0 U_{L0L1}+(1-\frac{\eta G}{g})L_1 U_{L1L1}+b_1\,F_1 gH\,U_{C1}-\frac{\eta G}{g}U_{L1}\}.$$

with $b_1 \equiv [\sum_{i=0}^{1}C_i U_{CiC1}+L_0 U_{L0C1}+(1-\frac{\eta G}{g})L_1 U_{L1C1}]/U_{C1}$

$\Leftrightarrow$

$$U_{L1}+F_1 gH\,U_{C1} = -a_1[\sum_{i=0}^{1}C_i U_{CiL1}+L_0 U_{L0L1}+(1-\frac{\eta G}{g})L_1 U_{L1L1}-(b_1+\frac{\eta G}{g})U_{L1}] \quad (33)$$

with $a_1 \equiv \tilde{\lambda}/[1+\tilde{\lambda}+b_1\,\tilde{\lambda}]$.

To make a clear case for differentiated taxation of qualified and nonqualified labour, evaluate (32) and (33) for the utility specification

$$U \equiv V(C_0,C_1)-\sum_{i=0}^{1}D(L_i) \quad (34)$$

with some linear homogenous function V. Then, $b_i=0$ (i=0,1) and

$$U_{L0}+F_0 H\,U_{C0} = \frac{\tilde{\lambda}}{1+\tilde{\lambda}}L_0 D_0'' \quad (32')$$

$$U_{L1}+F_1 gH\,U_{C1} = \frac{\tilde{\lambda}}{1+\tilde{\lambda}}[(1-\frac{\eta G}{g})L_1 D_1''-\frac{\eta G}{g}D_1']\,. \quad (33')$$

Dividing (33') through by (32') and setting $-F_0 H\frac{U_{C0}}{U_{L0}}-1\equiv\tau_0$, $-F_1 gH\frac{U_{C1}}{U_{L1}}-1\equiv\tau_1$ $\Leftrightarrow$ $F_i=(1+\tau_i)\omega_i$, $L_i D_i''/D_i'\equiv\nu_i$ gives us

$$\frac{\tau_1}{\tau_0} = \frac{(1-\frac{\eta G}{g})\nu_1 - \frac{\eta G}{g}}{\nu_0} \quad . \tag{35}$$

Interpret the left-hand side of (35) as the ratio of second-best optimal tax rates and ν_i as the inverse of labour supply in period i=0,1. For $\eta = 0$, (35) is the familiar *Inverse Elasticity Rule*. According to this rule wage taxes should be set inversely proportional to the wage elasticities of labour supplies $1/\nu_i$. This rule is extended by (35) to cope for endogenous education. The effect of education is to reduce the tax on qualified labour relative to the tax on nonqualified labour. Just note that $(1-\frac{\eta G}{g})\nu_1 - \frac{\eta G}{g} < \nu_1$. See Richter (2008) who derives a similar rule for the static framework. The deviation from the Inverse Elasticity Rule increases in $\eta G / g$. Hence the deviation increases in the elasticity of the learning function, $\eta = EG'/G$, and in the share, $G/g = G/(G+1-\delta_H)$, that newly formed human capital has in the periodic change in the stock of human capital.

Proposition 4: Assume selfish behaviour and U to satisfy (34). On a balanced growth path it is then second-best optimal to tax labour according to the Inverse Elasticity Rule (35). The effect of endogenous education is to reduce the tax on qualified labour relative to the tax on nonqualified labour.

The reader may want to learn how second-best policy translates into tax and subsidy rates. However, the finite-period Ramsey tax analysis only lends itself for an implicit determination of the policy rates. In the present context an explicit determination encounters even more difficulties than usual. Just for the sake of illustration, consider the special case in which the utility specification (34) and the Inverse Elasticity Rule (35) hold. Hence saving should remain untaxed. Denote by σ the rate by which the monetary cost of education should be subsidized in second best, $f = (1+\sigma)\varphi$. The optimal set of rates τ_1, τ_2 and σ must solve three equations. These are (35), the government budget constraint and the requirement of subsidizing education efficiently. The latter means that the extreme hand-sides of the following chain are equal:

$$F_0 L_0 + (F_0 + f)(\frac{g}{G'} - E) + \Delta = R\frac{F_0 + f}{G'} - F_1 L_1 = \frac{R}{G'}[\frac{\tau_0 - \tau_1}{1+\tau_0} F_0 + \frac{\sigma - \tau_1}{1+\sigma} f] \tag{36}$$

The first equality characterizes second-best efficient education policy. It is derived from (10) after correcting for the second-best efficient wedge Δ. The second equality characterizes the optimal private choice of education. It is derived from (15) and (14) after setting $F_i = (1+\tau_i)\omega_i$, $f = (1+\sigma)\varphi$ and after making some straightforward substitutions. The condition suggests that the need to subsidy the monetary cost of education and to tax nonqualified labour (relative to qualified labour) increases both in the second-best wedge Δ and in the intergenerational externality of human capital accumulation, $F_0 L_0 + (F_0 + f)(\frac{g}{G'} - E)$. Note however that this partial analytical interpretation ignores the fact that the efficient policy rates are jointly determined by (35), (36) and the government budget constraint.

4. Optimal taxation in the OLG model with altruistic individuals

The perfectly altruistic individual is assumed to maximize $\tilde{U}_t \equiv U(C_{0t}, C_{1t} L_{0t}, L_{1t}) + \beta \tilde{U}_{t+1}$ which by recursive substitution amounts to maximizing (3) in $C_{0t}, C_{1t} L_{0t}, L_{1t}, E_t, H_t, K_t$ (t=0,1, ..) subject to the human capital accumulation constraint (1) and the dynasty's budget constraint,

$$\sum_{t=0}^{\infty}[\pi_{t+1}\omega_{1t}L_{1t}H_t + \pi_t \omega_{0t}(L_{0t} - E_t)H_{t-1}]$$

$$= \sum_{t=0}^{\infty}[\pi_t C_{0t} + \pi_{t+1}C_{1t} + \pi_t \varphi_t E_t H_{t-1} + (\pi_t - R_{t+1}\pi_{t+1})K_t]. \qquad (\lambda) \qquad (37)$$

The first-order conditions are (t=0,1, ..):

$$\beta^t U_{C0t} = \lambda \pi_t, \ \ \beta^t U_{C1t} = \lambda \pi_{t+1}, \ \ \omega_{0t} H_{t-1} U_{C0t} = -U_{L0t}, \ \ \omega_{1t} H_t U_{C1t} = -U_{L1t}, \qquad (38)$$

$$\mu_t G_t' = (\varphi_t + \omega_{0t}) U_{C0t}, \quad R_{t+1} = \pi_t / \pi_{t+1} \qquad (39)$$

$$\lambda \pi_{t+1} [\omega_{1t} L_{1t} + \omega_{0t+1}(L_{0t+1} - E_{t+1}) - \varphi_{t+1} E_{t+1}] = \beta^t \mu_t - \beta^{t+1} [G_{t+1} + 1 - \delta_H] \mu_{t+1} \qquad (40)$$

The latter condition implies

$$\lambda \sum_{t=0}^{\infty} \pi_{t+1}[\omega_{1t} L_{1t} + \omega_{0t+1}(L_{0t+1} - E_{t+1}) - \varphi_{t+1} E_{t+1}] H_t \underset{(40)}{=} \sum_{t=0}^{\infty}[\beta^t \mu_t H_t - \beta^{t+1} \mu_{t+1} H_{t+1}]$$

$$= \mu_0 H_0 \underset{(39)}{=} \frac{\varphi_0 + \omega_{00}}{G_0^{'}} U_{C00} H_0 \tag{41}$$

Multiplying the budget constraint (37) through by λ and using (38), (39), and (41) to substitute for $\lambda\pi_t, \lambda\pi_{t+1}, \omega_{0t}, \omega_{1t}, R_{t+1}$ in (37) yields the Incentive Constraint

$$\sum_{t=0}^{\infty} \beta^t \sum_{i=0}^{1} C_{it} U_{Cit} = B \qquad (\tilde{\lambda}) \tag{42}$$

with

$$B \equiv \{[\omega_{00}(L_{00} - E_0) - \varphi_0 E_0] H_{-1} + \frac{\varphi_0 + \omega_{00}}{G_0^{'}} H_0\} U_{C00} .$$

Similarly, (38) and (39) can be used to substitute for $\lambda\pi_{t+1}, \omega_{0t+1}, \omega_{1t}, \mu_t$ in (40) which leaves us with (t=0,1, ..)

$$-L_{1t} U_{L1t} - \beta [(L_{0t+1} - E_{t+1}) U_{L0t+1} + \varphi_{t+1} E_{t+1} U_{C0t+1} H_t]$$

$$\underset{(40)}{=} \{\mu_t - \beta [G_{t+1} + 1 - \delta_H] \mu_{t+1}\} H_t = \mu_t H_t - \beta \mu_{t+1} H_{t+1}$$

$$\underset{(39)}{=} [\varphi_t U_{C0t} - U_{L0t} \frac{1}{H_{t-1}}] \frac{H_t}{G_t^{'}} - \beta [\varphi_{t+1} U_{C0t+1} - U_{L0t+1} \frac{1}{H_t}] \frac{H_{t+1}}{G_{t+1}^{'}} . \qquad (\gamma_t \beta^t) \tag{43}$$

The planner maximizes (3) in $C_{0t}, C_{1t} L_{0t}, L_{1t}, E_t, H_t, K_t, \varphi_t$ (t=0,1, ..) subject to the Resource Constraint (2), the accumulation constraint (1), and the behavioural constraints (42) and (43). Note however that φ_{t+1} (t=0,1, ..) only appears explicitly in condition (43). By contrast, the planner's objective function as well as the constraints (1), (2), and (42) are independent of φ_{t+1}. (43) can therefore be treated as a relationship defining the "free variable" φ_{t+1}. Hence the planner's problem is equivalent to the simplified version in which (3) is maximized in $C_{0t}, C_{1t} L_{0t}, L_{1t}, E_t, H_t, K_t$ (t=0,1, ..), and φ_0 subject to (1), (2), and (42). The same kind of solution procedure has been used before by Atkeson, Chari, and Kehoe (1999) and others before. We first study those first-order conditions of the simplified planner's problem which are associated with variables which do not enter the Incentive Constraint (42) or which drop out when making particular assumptions. The optimization with respect to those variables is not affected by (42) and should therefore remain undistorted.

Proposition 5: Assuming altruistic behaviour, then it is second-best efficient not to distort the accumulation of human capital for all generations (t=1,2, ..) except the first.

Proposition 6: Assuming altruistic behaviour, weak separability between consumption and non-leisure and assuming linear homogeneity of V, then it is second-best efficient not to distort the accumulation of capital for all generations (t=1,2, ...) except the first.

Proposition 7: Assuming altruistic behaviour, additive separability between consumption and non-leisure and assuming linear homogeneity of V, then it is second-best efficient to tax qualified and nonqualified labour uniformly for all generations (t=1,2, ..) except the first.

The proof of Proposition 5 is rather straightforward. Just note that the variables E_t, H_t, K_t (t=1,2, ..) do not enter the Incentive Constraint. Taking partial derivatives of the Lagrange function with respect to these variables and substituting for the Lagrange variables μ_t, α_t yields the efficiency condition (8) for t=1,2, ... The proof of Proposition 6 parallels the one of Proposition 1 and is therefore skipped. The proof of Proposition 7 is as follows. Set

$$W_t \equiv U_t + \tilde{\lambda} \sum_{i=0}^{1} C_{it} U_{Cit} \quad .$$

Additive separability of U implies $W_{Lit} = U_{Lit}$ (i=0,1). Taking partial derivatives of the Lagrange function with respect to K_t, L_{0t}, L_{1t}, yields (24) and

$$W_{L0t} + \alpha_t \; F_{0t} H_{t-1} = 0 = W_{L1t} + \alpha_{t+1}\beta \; F_{1t} H_t \quad (t=1,2, ..).$$

Taxes on labour are uniform if, and only if, $-F_{0t} H_{t-1} \dfrac{U_{C0t}}{U_{L0t}} - 1 \equiv \tau_{0t} = \tau_{1t} \equiv -F_{1t} H_t \dfrac{U_{C1t}}{U_{L1t}} - 1 \;\Leftrightarrow$ $\dfrac{U_{C0t}}{\alpha_t} = \dfrac{U_{C1t}}{\alpha_{t+1}\beta} \underset{(24)}{\Leftrightarrow} U_{C0t} / U_{C1t} = F_{Kt} + 1 - \delta_K$ (t=1,2, ..) which is the condition of efficient saving. The assumptions made for Proposition 7 are such that Proposition 6 applies. Hence saving should not be taxed and *a fortiori* labour should be taxed uniformly (t=1,2, ..).

With regard to the literature Proposition 6 can hardly surprise. This proposition fully stands in the Chamley-Judd tradition. Proposition 7 is less obvious and it even allows us to qualify the main result of Erosa and Gervais (2002) stating that it is generally optimal to differentiate labour taxes across the individual lifecycle. The intuitive explanation for this result is that labour supplied in the second life period is another commodity as labour supplied in the first period. However Erosa and Gervais assume selfish individuals and Proposition 4 confirms their result. The qualification suggested by Proposition 7 is that altruism removes the need to employ age-dependent labour taxes for descendent generations. Finally, Proposition 5 is interesting because it goes beyond the Chamley-Judd literature. It holds for arbitrary utility functions and without assuming balanced growth. I.e. Proposition 5 is logically stronger than Propositions 6 and 7. And it is also much stronger than the Education Efficiency Proposition obtained in static Ramsey analysis. In the finite-period framework it is only second-best efficient not to distort education if the earnings function is weakly separable in qualified labour supply and education and if the elasticity with respect to the latter is constant (Jacobs and Bovenberg, 2008; Bovenberg and Jacobs, 2005; Richter, 2006). By contrast, Proposition 5 holds for functions *G* which need not be isoelastic. This reminds one of the Production Efficiency Theorem of Diamond and Mirrlees (1971). According to this theorem the allocation of intermediate goods should not be distorted in second best given that no pure profits accrue to the private sector. This is just what holds in the present model. Investment in human capital is modelled as an intermediate good in the sense that it does not affect the Incentive Constraint for t=1,2, ... Furthermore, no pure profit accrues to the private sector. Just note that because of constant returns to scale in production the only pure profit is income earned by the parent generation of generation zero. This income equals $F_0 - (L_{00} - E_0)H_{-1}F_{00}$. It does not show up in the dynasty's budget constraint (37). Because of Walras Law it necessarily accrues to the government budget. The Production Efficiency Theorem is therefore applicable and Proposition 5 can be considered to be a corollary.

The government has to finance the exogenous cash flow of government expenditures A_t (t=0,1, ..). If the amount of pure profit earned by the government is insufficient, distortionary taxes have to be employed to balance the budget. In this case, the Incentive Constraint is binding and it cannot be ruled out that it is efficient to distort the choice of education of generation zero. This raises the question of how to design optimal human capital policy for generation zero. As we are going to learn, the answer comes close to what has been shown to be efficient in the world of selfish individuals. More precisely, a case is made for subsidizing the human capital investment of generation zero. To show this we maximize (3) subject to (1), (2), (42), and (43). Taking partial derivatives of the Lagrange function yields the following results after some straightforward manipulations have been made.

$$\frac{\partial}{\partial \varphi_0}: \quad \gamma_0 = -\tilde{\lambda}\,[1 - \frac{E_0 G_0^{'}}{H_0 / H_{-1}}] \tag{44}$$

$$\frac{\partial}{\partial \varphi_1}: \quad \gamma_1 = \gamma_0\,[1 - \frac{E_1 G_1^{'}}{H_1 / H_0}] \tag{45}$$

$$\frac{\partial}{\partial E_0}: \quad \frac{\mu_0}{\alpha_0} \underset{(44)}{=} \frac{f_0 + F_{00}}{G_0^{'}} - \frac{\tilde{\lambda}}{\alpha_0} U_{C00} \frac{\varphi_0 + \omega_{00}}{G_0^{'}} [1 + \frac{G_0^{''} E_0}{G_0^{'}}] \tag{46}$$

$$\frac{\partial}{\partial E_1}: \quad \frac{\mu_1}{\alpha_1} \underset{(44),(45)}{=} \frac{f_1 + F_{01}}{G_1^{'}} - \frac{\tilde{\lambda}}{\alpha_1} U_{C01} \frac{\varphi_1 + \omega_{01}}{G_1^{'}} [1 - \frac{E_0 G_0^{'}}{H_0 / H_{-1}}][1 + \frac{G_1^{''} E_1}{G_1^{'}}] \tag{47}$$

Making use of (44) – (47) yields

$$\frac{\partial}{\partial H_0}: \quad \Delta_0 = \frac{f_0 + F_{00}}{G_0^{'}} (F_{K0} + 1 - \delta_K) - \frac{f_1 + F_{01}}{G_1^{'}} (G_1 + 1 - \delta_H)$$

$$- F_{10} L_{10} - [F_{01} \cdot (L_{01} - E_1) - f_1 E_1] \tag{48}$$

with

$$\Delta_0 \equiv \frac{\tilde{\lambda}}{\alpha_0} U_{C00} \cdot MC_0^{HC} \cdot (F_{K0} + 1 - \delta_K)$$

$$- \frac{\tilde{\lambda}}{\alpha_1} U_{C01} \cdot MC_1^{HC} \cdot [1 - \frac{E_0 G_0^{'}}{H_0 / H_{-1}}](G_1 + 1 - \delta_H) \tag{49}$$

$$MC_t^{HC} \equiv \frac{\varphi_t + \omega_{0t}}{G_t^{'}} [1 + \frac{G_t^{''} E_t}{G_t^{'}} - \frac{G_t^{'} E_t}{H_t / H_{t-1}}] \quad (t=0,1). \tag{50}$$

The definitions of the symbols Δ_0 and MC_t^{HC} are set such that similarities with (29) and (18) are stressed. Without making additional assumptions, it is clearly difficult to sign Δ_0. In particular, negative values cannot be ruled out for appropriate choices of initial values. To neutralize the impact of initialization assume

$$MC_t^{HC} \text{ and } \frac{\tilde{\lambda}}{\alpha_t} U_{C0t} \text{ to be both positive and constant for } t\text{=0,1 and} \tag{51}$$

$$F_{K0} + 1 - \delta_K > G_1 + 1 - \delta_H . \tag{52}$$

This obviously implies positivity of Δ_0. Note that $\frac{E_0 G_0^{'}}{H_0 / H_{-1}} < \frac{E_0 G_0^{'}}{G_0} = \eta_0 < 1$.

Proposition 8: Assume (51) and (52). Then it is second-best efficient to subsidize the human capital investment of generation zero.

From a purely mathematical point of view Proposition 8 comes close to a triviality. The assumptions (51) and (52) are so strong that the result is obvious. The true significance of Proposition 8 is however a conceptual one. Proposition 8 parallels Proposition 3 and thus allows us to tell a unifying story for selfish and altruistic individuals. Altruism well reduces the need to subsidize education relative to *laissez-faire*. Altruism also implies that second-best tax policy for descendent generations is more like first-best policy. The accumulation of human capital should remain undistorted and labour taxes need not be differentiated across the individual life-cycle. The short-run policy recommendations for altruism however fully resemble the long-run recommendations for selfishness. Labour has to be taxed and the resulting decline in growth has to be compensated by subsidizing human capital accumulation relative to the first-best. Whether saving should be taxed is not a matter of selfishness or altruism. It primarily depends on assumptions made with regard to consumption preferences.

5. Summary

The accumulation of human capital may suffer from all sorts of potential inefficiencies. Most of them have simply been assumed away by the present study. Such a procedure is, no doubt, debatable. Critical is the ignoring of potential reasons of capital market or policy failure. Even more critical is the ignoring of individual heterogeneity and informational asymmetry. Still, the procedure is defended with the objective to study efficient taxation in Ramsey's tradition. More precisely, the paper aims at bridging the gap that separates the two strands of Ramsey tax analyses which exist for the finite and the infinite planning horizons. Our knowledge of efficient human capital policy in Ramsey's tradition is largely shaped by incompatible results derived for static and dynamic analysis. The results derived in dynamic analysis suggest that education should not be distorted in the long run just as saving should not be distorted in the long run. Hence it seems as if no difference should be made between human and nonhuman capital policies. By way of contrast, static analysis strongly suggests such differences. Whether education should be distorted or not appears to depend primarily on the question of how education affects the individual's earning potential. More precisely, only if the earnings function is weakly separable in qualified labour supply and education and if the elasticity with respect to the latter is constant, should the choice of education not be distorted by second-best policy (Jacobs and Bovenberg, 2008). On the other hand, whether saving should be distorted or not appears to depend on the taxpayer's consumption preferences. More precisely, saving should not be taxed if the taxpayer's utility is weakly separable between consumption and labour/nonleisure and homothetic in consumption (Atkinson and Stiglitz, 1972).

The model filling the gap between finite and infinite Ramsey tax analysis is one with overlapping generations. The present paper studies human capital policy in such a model with overlapping generations and endogenous growth. There have been earlier attempts to do the same. In view of the present study two attempts deserve to be cited more than others. These are the papers by Atkeson, Chari and Kehoe (1999) and Docquier, Paddison and Pestieau (2006). The most conspicuous differences to the present study are the following ones. The focus of the present study is on human capital accumulation while the focus of Atkeson et al. is on nonhuman capital. Their paper contains extensions to both endogenous education and overlapping generations but it fails to integrate the two. The paper by Docquier et al. is one integrating both, endogenous education and overlapping generations. However, the paper is none on endogenous labour supply and second-best taxation. The authors assume the availability of non-distortionary tax instruments which the present study does not. In a sense, the present paper starts where Atkeson et al. and Docquier et al. stop. It goes beyond Atkeson et al. by integrating endogenous education and overlapping generations and it goes beyond Docquier et al. by endogenizing labour supply and by doing second-best tax analysis.

The present paper studies two reasons of potential allocational inefficiency. One is the lacking availability of non-distortionary tax instruments. The other is individual selfishness. Taxpayers are assumed to externalize the positive effect that their human capital investments have on the productivity of descendent generations. As stressed by Docquier et al. selfishness is the source of an intergenerational externality. It gives reason to subsidize human capital investments relative to *laissez-faire*. Such subsidization however requires government revenues. In the framework studied by Docquier et al. it is efficient to subsidize human capital accumulation up to the first-best level where marginal social costs equal marginal social returns. The result assumes the availability of non-distortionary tax instruments. The key assumption of the present study however is that no tax instruments are available that would allow the government to raise the revenue needed to subsidize education without creating distortions. It then turns out to be efficient at a balanced growth path to subsidize human capital accumulation even beyond the first-best level. In other words, it is efficient in the long run to combine positive tax wedges in the labour market with a negative tax wedge for human capital investment. First-best incentives for accumulating human capital would be too weak to compensate efficiently for the decline in long-run growth induced by labour taxation. The strength of efficient positive distortion is shown to increase in (i) the cost resulting from the non-availability of lump-sum taxes, (ii) the dynamic cost of education, and (iii) the growth gap. Furthermore, it turns out to be efficient to tax labour such that qualified labour is less distorted than nonqualified labour.

If taxpayers are altruists with respect to descendent generations, a clear reason for government intervention disappears. The effect that education has on descendent generations' productivity is internalized by altruists. The only remaining inefficiency modelled in this paper is caused by the need to employ distortionary taxes for financing exogenous government expenditures. As it turns out all generations except the first one should still be given non-distorted incentives for accumulating human and nonhuman capital. Furthermore, labour should be taxed uniformly across the individual life-cycle. The latter allows us to qualify the main result of Erosa and Gervais (2002) who stress the need to employ age-dependent labour taxes in second best. Obviously, altruism removes the need. In view of the Chamley-Judd literature results suggesting non-distortionary taxation may not be too surprising. Striking is however the strength of the result concerning human capital accumulation. While the other results on non-distortionary taxation require special utility functions, the result on human capital accumulation holds without any comparable qualification. One only has to assume that no pure profit accrues to the private sector. It is argued that this result on efficient education policy is best interpreted as a corollary to the Production Efficiency Theorem of Diamond and Mirrlees (1971).

The results on non-distortionary taxation do not require removing all possible distortions. On the contrary, the labour supply of descendent generations will be distorted if the government has to finance exogenous government expenditures by relying on distortionary instruments. The results on non-distortionary taxation do neither extend to the dynasty's first generation indexed by zero in the present paper. A more specific characterization of optimal policy for generation zero is difficult as the specific features strongly depend on initial values. After neutralizing the impact of initialization a case can however be made for subsidizing education relative to the first best. The reason is the same identified for the scenario with selfish individuals. Taxing labour has a negative effect on education and growth and education policy has to compensate for this dynamic effect. The unifying bottom line for selfish and altruistic individuals is as follows. Altruism well reduces the need to subsidize education relative to *laissez-faire* and altruism also implies that descendent generations should be given non-distorted incentives for accumulating human capital. The short-run policy recommendations for altruism however agree with the long-run recommendations for selfishness. Labour has to be taxed and the resulting decline in growth has to be compensated by subsidizing education relative to the first-best. Whether saving should be taxed is not a matter of selfishness or altruism. It primarily depends on assumptions made with regard to consumption preferences.

References

Atkeson, A., V.V. Chari, and P.J. Kehoe, 1999, Taxing capital income: A bad idea. Federal Reserve Bank of Minneapolis Quarterly Review 23, 3−17.

Atkinson, A.B. and J.E. Stiglitz, 1972, The structure of indirect taxation and economic efficiency. Journal of Public Economics 1, 97-119.

Bovenberg, A.L. and B. Jacobs, 2005, Redistribution and learning subsidies are Siamese twins. Journal of Public Economics 89, 2005–2035.

Bull, N., 1993, When all the optimal dynamic taxes are zero. Federal Reserve Board Working Paper 137.

Chamley, C., 1986, Optimal taxation of capital income in general equilibrium with infinite lives. Econometrica 54, 607−622.

Diamond, P. and J.A. Mirrlees, 1971, Optimal taxation and public production, I Production efficiency, II Tax rules. American Economic Review 61, 8–27, 261–278.

Docquier, F., O. Paddison, and P. Pestieau, 2006, Optimal accumulation in an endogenous growth setting with human capital. IZA Discussion Paper No. 2081.

Erosa, A. and M. Gervais, 2002, Optimal taxation in life-cycle economies. Journal of Economic Theory 105, 338-369.

Jacobs, B. and A.L. Bovenberg, 2008, Optimal taxation of human capital and the earnings function. CESifo Working Paper No. 2250.

Jones, L.E., R.E. Manuelli, and P.E. Rossi, 1993, Optimal taxation in models of endogenous growth. Journal of Political Economy 101, 485-517.

Jones, L.E., R.E. Manuelli, and P.E. Rossi, 1997, On the optimal taxation of capital income. Journal of Economic Theory 73, 93−117.

Judd, K.L., 1985, Redistributive taxation in a simple perfect foresight model. Journal of Public Economics 28, 59−83.

Lipsey, R.G., and K. Lancaster, 1956/57, The general theory of second best. Review of Economic Studies 24, 11-32.

Ramsey, F.P., 1927, A contribution to the theory of taxation. Economic Journal 37, 47-61.

Richter, W.F., 2007, Taxing human capital efficiently – The double dividend of taxing nonqualified labour more heavily than qualified labour. Ruhr Economic Papers 12.

Richter, W.F., 2008, Taxing education in Ramsey's tradition. CESifo Working Paper No. 2586.

Sandmo, A., 1974, A note on the structure of optimal taxation. American Economic Review 64, 701-6.