ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Richter, Wolfram F.; Braun, Christoph

Working Paper

Efficient subsidization of human capital accumulation with overlapping generations and endogenous growth

IZA Discussion Papers, No. 4629

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Richter, Wolfram F.; Braun, Christoph (2009) : Efficient subsidization of human capital accumulation with overlapping generations and endogenous growth, IZA Discussion Papers, No. 4629, Institute for the Study of Labor (IZA), Bonn

This Version is available at:
https://hdl.handle.net/10419/35956

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

DISCUSSION PAPER SERIES

IZA DP No. 4629

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

Wolfram F. Richter
Christoph Braun

December 2009

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

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

Wolfram F. Richter
TU Dortmund University
and IZA

Christoph Braun
Ruhr Graduate School in Economics
and TU Dortmund University

Discussion Paper No. 4629
December 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA Discussion Paper No. 4629
December 2009

ABSTRACT

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

This paper studies second-best policies in an OLG model in which endogenous growth results from human capital accumulation. When young, individuals decide on education, saving, and nonqualified labour. When old, individuals supply qualified labour. Growth equilibria are inefficient in *laissez-faire* because of distortionary taxation. The inefficiency is exacerbated if selfish individuals externalize the positive effect of education on descendents' productivity. It is shown to be second best not to distort education if the human capital investment function is isoelastic. If the function is not isoelastic, a case is made for subsidizing education even relative to the first best.

JEL Classification: H21, I28, J24

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

Corresponding author:

Wolfram F. Richter
TU Dortmund University
Department of Economics
44221 Dortmund
Germany
E-mail: Wolfram.Richter@tu-dortmund.de

* This paper was prepared for the Conference in Honour of Robin Boadway, Queen's University, Kingston, Ontario, 14-15 May 2009. The authors thank their discussants, Maria Racionero and Jim Davies, and also Lutz Arnold, Berthold Wigger, and two anonymous referees for helpful suggestions.

1. Introduction

The implications of adding endogenous education to the Ramsey problem of efficient taxation are by now well understood. Unfortunately, they strongly depend on whether the representative taxpayer plans for finite or infinite periods. If the taxpayer's planning horizon is infinite, the rationale for employing distortionary linear taxes turns out to be weak. This point was originally made by Chamley (1986) and Judd (1985) with respect to capital taxes. It extends, however, to the model with endogenous education, as has been demonstrated by Bull (1993), Jones, Manuelli, and Rossi (1993, 1997), and Atkeson, Chari, and Kehoe (1999). The question whether human or nonhuman capital is accumulated is largely irrelevant. In the long run neither accumulation should be slowed down by distortionary taxation.

Finite Ramsey tax analysis, by contrast, suggests treating human and nonhuman capital differently. If the taxpayer's planning horizon is finite, it depends on consumption preferences whether taxing saving is efficient or not. In particular, saving should be untaxed only if the taxpayer's utility is weakly separable between consumption and labour and homothetic in consumption (Atkinson and Stiglitz, 1972; Sandmo, 1974). By contrast, the design of efficient education policy is more a reflection of the specific properties of the earnings function. This function has to be weakly separable in qualified labour supply and education and the elasticity with respect to the latter has to be constant if it shall be second best not to distort the choice of education. To ensure an undistorted choice of education, the effective return must equal the effective cost before taxes and subsidies (Jacobs and Bovenberg, 2008; Bovenberg and Jacobs, 2005; Richter, 2007). Furthermore, it is second best to tax labour in such a way that qualified labour is distorted less than nonqualified labour (Richter, 2009).

Studying optimal tax design in Ramsey's tradition is known to suffer from various shortcomings. The most obvious ones are its focussing on a representative taxpayer and its ignoring heterogeneity and informational asymmetry. Critical is also that such study ignores possible causes of capital market or policy failures. The present paper, however, puts aside all such objections. Its sole objective is to contribute to the literature that tries to bridge the gap between the finite and the infinite Ramsey policy analyses. More specifically, the paper explores policy implications for education, labour, and saving in a framework in which human capital accumulation is modelled as a dynamic activity affecting the well-being of descendent generations.

The model chosen is one with overlapping generations and endogenous growth. Individuals live for two periods. They decide on education, saving, and nonqualified labour in their youth.

They supply qualified labour when old. The productivity of qualified labour increases in the stock of human capital inherited from preceding generations, and it also increases in own educational investments. Individuals either may be perfect altruists with respect to descendent generations or may behave selfishly. The implications of selfishness have been studied before by Wigger (2002, Sec. 3.4) and Docquier et al. (2007) for a framework in which the government is not constrained in the use of policy instruments. It is shown that decentralizing the first best requires subsidizing education. The present paper goes beyond these earlier studies by endogenizing labour supply and by assuming that the government can only employ linear policy instruments. Most remarkably, major results characterizing efficient static policy extend to the dynamic framework. For example, it is shown to be second best to distort qualified labour less than nonqualified labour. In particular, it is second best not to distort education if the human capital investment function is isoelastic in education. It is argued, however, that such constant elasticity has debatable implications in a dynamic framework. It implies that the human capital stock accumulated by preceding generations melts down to zero if just one generation stops investing. More appealing is the assumption that the elasticity of the investment function is increasing and that the human capital stock does not depreciate completely if just one generation fails to invest. If this is the case, it is second best at balanced growth to subsidize education even in relation 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. Not surprising is the need to subsidize education relative to *laissez faire*. This is so because the intergenerational externalities of human capital investments have to be internalized.[2] *A priori* it is not obvious, however, why investments should even exceed the first-best level. 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 education if labour has to be taxed and if the level of comparison is the first best. The contrary, however, is true. The key assumption is the increasing elasticity of the human capital investment function with respect to education. The effect is that it is second best to subsidize education in static analysis, and this effect is shown to extend to the dynamic framework. At balanced growth the need to subsidize increases in the derivative of the investment function's elasticity and in two further factors. One factor is the Lagrange multiplier on the planner's implementability constraint, and the other is the gap between the marginal return to capital and the rate of balanced growth. In other words, the

[2] The need is highlighted by various earlier studies. An example is Del Rey and Racionero (2002).

more binding the non-availability of lump-sum taxes is and the more deficient the growth is, the more should human capital accumulation be encouraged beyond what appears to be first best.

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. That second source is the need to employ distortionary taxes for financing 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. The accumulation of human capital should not be distorted, and this result is obtained for arbitrary utility and human capital investment functions. Furthermore, qualified and nonqualified labour should be taxed uniformly across the life cycle if, and only if, saving should not be distorted. Such results strongly contrast with those derived for the case of selfish individuals.

The results obtained for the first generation are less contrasting. In particular, it is second best not to distort the first generation's educational choice if the human capital investment function is isoelastic in education. If, however, this function fails to be isoelastic, the optimal education policy for the first generation depends on initial values. On neutralizing the effect of initialization by assuming balanced growth and assuming an increasing elasticity of the human capital investment function, it turns out to be second best to subsidize education. The reason is the same encountered when individuals are selfish. Increasing elasticity is the reason why it is second best to subsidize education in static analysis. This effect extends to the dynamic framework. The need to subsidize is the stronger the larger the derivative of the investment function's elasticity is, the more binding the non-availability of lump-sum taxes is, and the more deficient growth is.

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 invest in human capital. The short-run policy recommendations for altruism, however, agree with the long-run recommendations for selfishness. Labour has to be taxed, and – given that the elasticity of the human capital investment function is increasing – education should be subsidized relative to the first best. Whether saving should be taxed or not primarily depends on assumptions made with regard to consumption preferences.

There are countries in which the accumulation of human capital is effectively subsidized relative to the first best. According to the empirical evidence provided by Collins and Davies (2004) Canada is an example.[3] Trostel (1996) is able to demonstrate the efficiency-enhancing potential of subsidization by calibrating a Ramsey-type model and by carrying out sensitivity analysis. What has, however, been lacking so far is a normatively convincing justification.

The paper is structured as follows. Section 2 sets up the 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 who are altruistic towards descendent generations. Section 5 summarizes.

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

Consider a sequence of overlapping generations with individuals living for two periods. The index t refers to the generation and to the period in which the representative individual of generation t is young and in her life period 0. Lifetime utility is given by $U_t \equiv U(C_{0t}, C_{1t}, L_{0t}, L_{1t})$ with the arguments C_{0t}, C_{1t}, L_{0t}, and L_{1t} denoting consumption and non-leisure in the life periods i=0,1. Utility is strictly increasing in consumption, is strictly decreasing in non-leisure, and is strictly quasi-concave. Non-leisure in the second life period, L_{1t}, equals *qualified* labour supplied to the market in period t+1. By contrast, non-leisure in the first life period has to be divided between *nonqualified* labour supply $L_{0t} - E_t$ and education E_t. The effect of education is to increase human capital and labour productivity. H_{t-1} is the stock of human capital determining the productivity in period t. It is built up by generation $t-1$ and inherited by generation t. By spending time E_t on education, generation t

[3] OECD data provide further evidence of countries effectively subsidizing education relative to the first best. Unfortunately, however, the evidence strongly varies with the method used in estimating the private and the social net returns to education. There is the OECD study "Education at a Glance" of 2005 compiling internal rates of return for an individual obtaining a university-level degree. As one can infer from Tables A9.6 and A9.10 the vast majority of ten countries effectively subsidizes education relative to the first best. Opposing evidence is provided by OECD (2009, Indicators A8). This study compiles net present values from which one can infer that only a minority of 21 countries subsidizes effectively. Apparently, a closer look at the collection of the data is needed if such contradictory evidence is to be resolved. Doing so must however remain the object of future research.

determines the stock of human capital H_t effective in the second life period. The *human capital accumulation equation* is

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

The investment function $G_t \equiv G(E_t)$ is non-negative and strictly monotone increasing with elasticity $\eta \equiv EG'/G$ smaller than one. $\mu_t \beta^t$ is a Lagrange multiplier associated with the planner's problem we are about to set up. The functional specification (1) is standard in the endogenous growth literature. It can be traced back to Uzawa (1965), and it has been used since by Lucas (1988), Atkeson et al. (1999), Wigger (2002), and others. A key feature is that H_t is linear homogenous in H_{t-1}. Docquier et al. (2007) deviate by assuming H_t to be linear homogeneous in H_{t-1} and E_t. A notable implication of (1) is that time spent on education (*learning*) is the only variable input in the production of human capital. In particular, learning cannot be substituted by physical inputs. Such physical goods are used instead in fixed proportion with education. For simplicity's sake, the resulting direct costs are modelled as a linear function of education and inherited human capital, $f E_t H_{t-1}$, where the (*fee*) parameter f is exogenous.

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 is not productive before the second life period, and it depreciates at the rate δ_K. Production *F* is linear homogenous in capital and effective labour. The *resource constraint* is

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

with $F_t \equiv F(K_{t-1}, (L_{0t} - E_t)H_{t-1}, L_{1t-1}H_{t-1})$.

The variable A_t denotes exogenous government spending. Such spending may be of consumptive and/or productive use. As A_t is exogenous, we refrain from making it an explicit argument of the utility and/or production functions. When taking partial derivatives use is made of the following short forms:

$$F_{Kt} \equiv \frac{\partial F}{\partial K_{t-1}}, \qquad F_{L_0 t} \equiv \frac{\partial F}{\partial((L_{0t} - E_t)H_{t-1})}, \qquad F_{L_1 t} \equiv \frac{\partial F}{\partial(L_{1t-1}H_{t-1})}.$$

Qualified and nonqualified labour may be perfect substitutes in production, but they need not be. Human capital is obviously labour augmenting. Note that education incurs two kinds of

cost. There are *costs of forgone earnings*, $F_{L_0 t}\, E_t H_{t-1}$, besides (*monetary*) *direct costs*, $f\, E_t H_{t-1}$.

The planner maximizes

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

in $C_{0t}, C_{1t}, L_{0t}, L_{1t}, E_t, H_t$, and K_t (t=0,1,...) subject to the human capital accumulation equation (1) and the resource constraint (2). The parameters $K_{-1}, H_{-1}, L_{1-1} = L_{1,t=-1}$ are exogenously given. 0< β <1 is a discount factor. Assume that this maximization – like all others still to follow – is well behaved and that it has an interior solution for which all choice variables are strictly positive. We abstain from stating all the assumptions needed to guarantee a well-behaved maximization with interior solutions. Identifying those assumptions must remain the object of independent research efforts. In the present paper we just state those assumptions explicitly needed to derive meaningful first-order conditions of second-best policies. We study neither second-order conditions nor questions of existence. As argued in Richter (2009) and as will become clearer below, a well-behaved maximization requires a specification of $U_t = U(C_{0t}, C_{1t} L_{0t}, L_{1t})$ which is sufficiently concave to compensate for the lack of concavity of the human capital accumulation equation (1). The first-order conditions of the planner's maximization are as follows:

$$U_{C_0 t} = \alpha_t\,, \quad U_{C_1 t} = \alpha_{t+1}\beta\,, \quad F_{L_0 t}\, H_{t-1}\, U_{C_0 t} = -U_{L_0 t}\,, \quad F_{L_1 t+1} H_t\, U_{C_1 t} = -U_{L_1 t}\,, \tag{4}$$

$$F_{Kt+1} + 1 - \delta_K \;=\; U_{C_0 t} / U_{C_1 t} \;=\; U_{C_0 t} / \beta\, U_{C_0 t+1}\,, \tag{5}$$

$$\mu_t G_t^{'} = \alpha_t\, (f + F_{L_0 t})\,, \tag{6}$$

$$\alpha_{t+1}\beta\, [\, F_{L_1 t+1} L_{1t} + F_{L_0 t+1} \cdot (L_{0t+1} - E_{t+1}) - f\, E_{t+1}\,] = \mu_t - \beta\, G_{t+1}\, \mu_{t+1}\,. \tag{7}$$

The conditions (4) characterize efficient consumption and labour choices. The condition (5) characterizes efficient saving and efficient capital. The 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*,

$$F_{L_1 t+1} L_{1t} + F_{L_0 t+1}\, L_{0t+1} - (\, F_{L_0 t+1} + f\,)\, E_{t+1}$$

$$= \; [\, F_{Kt+1} + 1 - \delta_K \,] \frac{f + F_{L_0 t}}{G_t^{'}} \; - \; G_{t+1} \frac{f + F_{L_0 t+1}}{G_{t+1}^{'}} . \qquad (8)$$

For the sake of brevity we also speak of *efficient education* if (8) holds. The first term on the left-hand side, $F_{L_1 t+1} L_{1t}$, is the return to human capital accruing to generation *t* in the second life period, and the difference $F_{L_0 t+1}\, L_{0t+1} - (\, F_{L_0 t+1} + f\,)\, E_{t+1}$ is the return accruing to individuals of the next generation in their first life period. The factor

$$\frac{f + F_{L_0 t}}{G_t^{'}} \; = \; (\, f + F_{L_0 t}) H_{t-1} \frac{dE_t}{dH_t}$$

is the cost of human capital in period *t*, and $\frac{f + F_{L_0 t+1}}{G_{t+1}^{'}}$ is the cost of human capital one period later. Hence the right-hand side of (8) captures the cost resulting from investing in period *t* instead of postponing the investment to the next period. By separating terms referring to generation *t* from terms referring to generation *t*+1, we obtain from the efficiency condition (8)

$$F_{L_1 t+1} L_{1t} - [\, F_{Kt+1} + 1 - \delta_K \,] \frac{f + F_{L_0 t}}{G_t^{'}} \; = \; - \; F_{L_0 t+1}\, L_{0t+1} - (\, F_{L_0 t+1} + f\,)\, E_{t+1} [\, \frac{1}{\eta_{t+1}} - 1] \, . \qquad (9)$$

Because $\eta_{t+1} < 1$, the right-hand side of (9) is negative. Hence first-best policy requires generation *t*'s cost, $[\, F_{Kt+1} + 1 - \delta_K \,] \frac{f + F_{L_0 t}}{G_t^{'}}$, to exceed generation *t*'s return to human capital, $F_{L_1 t+1} L_{1t}$. The difference is the positive external effect on generation *t*+1, which has to be internalized by first-best policy when individuals are selfish.

We speak of *balanced growth* if the non-leisure choices $L_{0t} = L_0, L_{1t} = L_1$, and $E_t = E$ are constant across time while consumption, output, and both types of capital all grow at the common gross rate $G = G(E)$, so that we have $H_{t-1} = G^t H_{-1}, K_{t-1} = G^t K_{-1}, C_{it} = G^t C_{i0}$. At balanced growth, $F_{Kt+1} = F_K$ is constant across time. Assuming utility *U* to be homogeneous of degree *d* in consumption, we obtain $U_{C_0 t} = G^{(d-1)t}\, U_{C_0 0}$. In the case of homogeneity, the conditions (4) and (5) imply $F_K + 1 - \delta_K \; = \; G^{1-d} / \beta$. Furthermore, the *condition of*

transversality, $\beta^t U_{C_0 t} K_t \rightarrow 0$ for $t \rightarrow \infty$, implies $(\beta G^d)^t U_{C_0 0} G K_{-1} \rightarrow 0$ for $t \rightarrow \infty$, i.e., $\beta G^d < 1$. As a result, the return to capital exceeds the growth rate:

$$F_K + 1 - \delta_K = G^{1-d} / \beta > G. \tag{10}$$

The following analysis studies second-best policy with regard to education, to saving, and also to labour. The focal question, however, is whether it is second best to provide or not to provide efficient incentives for education. As we shall see, much depends on the functional specification of $G(E)$ and on 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 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 $C_{0t}, C_{1t}, L_{0t}, L_{1t}, E_t$, and K_t subject to the lifetime budget constraint

$$\begin{aligned} &\pi_t \omega_{0t} (L_{0t} - E_t) H_{t-1} + \pi_{t+1} \omega_{1t} L_{1t} G(E_t) H_{t-1} \\ &\qquad = \pi_t C_{0t} + \pi_{t+1} C_{1t} + \pi_t \varphi_t E_t H_{t-1} + (\pi_t - R_{t+1} \pi_{t+1}) K_t . \qquad (\lambda_t) \end{aligned} \tag{12}$$

In this optimization H_{t-1} is treated as an exogenous parameter. π_t is the price of consumption, ω_{0t} is the wage rate of nonqualified labour, ω_{1t} is the wage rate of qualified labour, φ_t is the direct (unit) cost of education, and R_{t+1} is the real return earned on savings, $\pi_t K_t$. All these prices and costs are after tax and subsidy. For each t there are six first-order conditions

$$U_{C_0 t} = \pi_t \lambda_t , \quad U_{C_1 t} = \pi_{t+1} \lambda_t , \tag{13}$$

$$\omega_{0t} \; H_{t-1} \, U_{C_0 t} = -U_{L_0 t} , \quad \omega_{1t} \; G(E_t) H_{t-1} \, U_{C_1 t} = -U_{L_1 t} , \tag{14}$$

$$\omega_{1t} L_{1t}\, G_t^{'}\, U_{C_1 t} = (\varphi_t + \omega_{0t})\, U_{C_0 t}\,, \quad R_{t+1} = \pi_t / \pi_{t+1}\,, \tag{15}$$

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

$$\sum_{i=0}^{1} [C_{it} U_{C_i t} + L_{it} U_{L_i t}] \;=\; \eta_t\, L_{1t} U_{L_1 t}\,. \qquad (\tilde{\lambda}_t \beta^t) \tag{16}$$

The condition (16) will assume the role of an *implementability constraint* in the planner's second-best problem. Because

$$-\eta_t \frac{L_{1t} U_{L_1 t}}{U_{C_0 t}} = (\varphi_t + \omega_{0t})\; E_t H_{t-1}\,,$$

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

$$MC_t^{HC} \equiv -\frac{d}{dH_t}[\,\eta_t \frac{L_{1t} U_{L_1 t}}{U_{C_0 t}}\,] = -\frac{L_{1t} U_{L_1 t}}{U_{C_0 t}} \frac{dE_t}{dH_t} \frac{d}{dE_t} \eta(E_t) \;=\; \frac{\varphi_t + \omega_{0t}}{G_t^{'}} \frac{E_t \eta_t^{'}}{\eta_t}\,. \tag{17}$$

By calling the cost dynamic we are not suggesting that the cost can be ignored in static analysis. Rather, we call it dynamic because it depends on the second-order elasticity of $G(E_t)$. The cost is increasing in the derivative of the elasticity of $G(E_t)$. If the elasticity $\eta_t = \eta(E_t)$ is constant, MC_t^{HC}=0 results. A critical implication of constant elasticity is $G(0)$=0. Because $H_t = G(E_t)\, H_{t-1}$, the stock of human capital built up by the preceding generations, H_{t-1}, then melts down to zero if generation *t* does not engage in education. This implies depreciation of one hundred percent.

This may be considered a quite extreme implication, and a specification of G with $G(0)=1-\delta_H>0$ may be considered to be more appealing. The latter means that some human capital is passed on to the next generation even if there are no new investments. To allow for this possibility, consider the case where $G(E)$ not itself but $G(E)-(1-\delta_H)$ is of constant elasticity η and where the rate of depreciation δ_H is less than one. In this case,

$$MC_t^{HC} = -\frac{L_{1t}U_{L_1t}}{U_{C_0t}}\frac{1}{H_t}\eta\frac{1-\delta_H}{G_t} > 0$$

is obtained. In what follows, MC_t^{HC} is assumed to be nonnegative, which according to (17) requires the elasticity of G to be weakly increasing.

Remark 1: Assume $G(E)-(1-\delta_H)$ to be isoelastic. The dynamic cost of education is positive if δ_H <1, and it is zero if δ_H =1.

The planner maximizes the sum of discounted lifetime utilities (3) in $C_{0t}, C_{1t}, L_{0t}, L_{1t}, E_t, H_t$, and K_t (t=0,1,...) subject to the implementability constraint (16), the human capital accumulation equation (1), and the resource constraint (2). The solutions are second best in the sense that they have to fulfil the implementability constraint 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 $\omega_{0t}, \omega_{1t}, \varphi_t$, and R_{t+1}. This means in particular that labour income can be taxed at different rates over an individual's life cycle. 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

$$W_t \equiv U_t + \tilde{\lambda}_t \{ \sum_{i=0}^{1} [C_{it}U_{C_it} + L_{it}U_{L_it}] - \eta_t \, L_{1t}U_{L_1t} \} \, . \tag{18}$$

The first-order conditions are as follows:

$$\frac{\partial}{\partial C_{0t}}, \frac{\partial}{\partial L_{0t}}: \quad W_{C_0t} = \alpha_t = -\frac{W_{L_0t}}{F_{L_0t}H_{t-1}}, \tag{19}$$

$$\frac{\partial}{\partial C_{1t}}, \frac{\partial}{\partial L_{1t}}: \quad W_{C_1t} = \alpha_{t+1}\,\beta = -\frac{W_{L_1t}}{F_{L_1t+1}H_t}, \tag{20}$$

$$\frac{\partial}{\partial K_t}: \quad \alpha_{t+1}\beta\,[\,F_{Kt+1}+1-\delta_K\,] = \alpha_t\,, \tag{21}$$

$$\frac{\partial}{\partial E_t}: \quad \mu_t\, G_t^{'}\, H_{t-1} = \tilde{\lambda}_t\, \eta_t^{'}\, L_{1t} U_{L_1 t} + \alpha_t\,(\,f + F_{L_0 t}\,)\, H_{t-1} \Rightarrow$$

$$\frac{\mu_t}{\alpha_t} \underset{(17)}{=} \frac{f + F_{L_0 t}}{G_t^{'}} - \frac{\tilde{\lambda}_t}{\alpha_t}\, U_{C_0 t}\, MC_t^{HC}\,, \tag{22}$$

$$\frac{\partial}{\partial H_t}: \quad \alpha_{t+1}\,\beta\,[\,F_{L_1 t+1} L_{1t} + F_{L_0 t+1}\cdot(\,L_{0t+1} - E_{t+1}\,) - \; f\,E_{t+1}\,] + \mu_{t+1}\,\beta\, G_{t+1} = \mu_t\,. \tag{23}$$

We shall derive characterizations of second-best policy with regard to saving, education, and labour.

We start with saving. As has been shown by Atkinson and Stiglitz (1972), Sandmo (1974), Atkeson, Chari and Kehoe (1999), and others, it is efficient not to distort saving if utility is weakly separable between consumption and non-leisure and is homothetic in consumption. This well-known result extends to the present framework. One only has to assume that utility can be written in the form

$$U = U(V(C_0, C_1), L_0, L_1) \tag{24}$$

with a linear homogeneous function *V*. In this case, the private and the social marginal rates of substitution in consumption are equal,

$$\frac{W_{C_0 t}}{W_{C_1 t}} = \frac{U_{C_0 t}}{U_{C_1 t}}\,. \tag{25}$$

The proof of this equality is straightforward. It suffices to prove that $W_{C_i t}\,/\,U_{C_i t}$ is constant in *i*=0,1. For the sake of simplicity the index *t* is suppressed:

$$\frac{W_{C_i}}{U_{C_i}} = \frac{1}{U_{C_i}}\,\frac{d}{dC_i}\,\{U + \tilde{\lambda}\,\sum_{i=0}^{1}[C_i U_{C_i} + L_i U_{L_i}\,] - \tilde{\lambda}\,\eta\, L_1 U_{L_1}\,\}$$

$$= 1 + \tilde{\lambda}\,\{1 + \sum_{j=0}^{1}[C_j\,\frac{U_{C_j C_i}}{U_{C_i}} + L_j\,\frac{U_{L_j C_i}}{U_{C_i}}\,] - \eta\, L_1\,\frac{U_{L_1 C_i}}{U_{C_i}}\,\}$$

$$= 1 + \tilde{\lambda}\,\{1 + V\,\frac{U_{VV}}{U_V} + \sum_{j=0}^{1} L_j\,\frac{U_{VL_j}}{U_V} - \eta\, L_1\,\frac{U_{VL_1}}{U_V}\,\} \;=\; \text{constant in } i\text{=0,1.}$$

If the equality (25) holds, it turns out to be optimal from the planner's perspective to equate the marginal rate of return to capital with the private marginal rate of substitution in consumption. This follows from (25) on making use of (19)–(21):

$$F_{Kt+1} + 1 - \delta_K = \frac{\alpha_t}{\alpha_{t+1}\beta} = \frac{W_{C_0 t}}{W_{C_1 t}} = \frac{U_{C_0 t}}{U_{C_1 t}} \ . \tag{26}$$

Proposition 1: If behaviour is selfish and if utility is weakly separable between consumption and non-leisure and homothetic in consumption, then it is second best not to distort saving.

We turn next to education. We first prove that it is efficient not to distort human capital accumulation if the investment function G is isoelastic. We do so by relying on (21)–(23), which are the first-order conditions with respect to K_t, E_t, and H_t. By making use of (21) and (22), (23) can be written as

$$[\, F_{L_1 t+1} L_{1t} + F_{L_0 t+1} \cdot (\, L_{0t+1} - E_{t+1}) - f\, E_{t+1}] + [\, \frac{f + F_{L_0 t+1}}{G'_{t+1}} - \frac{\tilde{\lambda}_{t+1}}{\alpha_{t+1}} U_{C_0 t+1}\, MC^{HC}_{t+1}\,] G_{t+1}$$

$$= [\, \frac{f + F_{L_0 t}}{G'_t} - \frac{\tilde{\lambda}_t}{\alpha_t} U_{C_0 t}\, MC^{HC}_t\,][\, F_{Kt+1} + 1 - \delta_K\,]. \tag{27}$$

Obviously, (27) equals (8) whenever $MC^{HC}_{t+1} = MC^{HC}_t = 0$. Remark 1 therefore yields

Proposition 2: Assume selfish behaviour. It is second best not to distort education if the human capital investment function $G(E)$ is isoelastic.

Proposition 2 is a dynamic version of the *education efficiency proposition*, well known from static tax analysis (Jacobs and Bovenberg, 2008; Bovenberg and Jacobs, 2005; Richter, 2007). An intuitive explanation is the following. The planner cares about two objectives. One objective is to minimize the efficiency loss resulting from distorted choices of consumption and non-leisure. The other objective is to minimize losses in the rent income generated by education. In general, these two minimizations are not separable, so that the planner has to trade off. Separability is only ensured if the human capital investment function is isoelastic. If

this is the case and if the set of policy instruments is sufficiently rich, it is efficient not to distort education and to minimize the efficiency loss resulting from distorted choices of consumption and non-leisure. According to Proposition 2 this result extends to the dynamic framework. It does not hold if the dynamic cost of education is positive.

To study this case set

$$\Delta_t \equiv \frac{\tilde{\lambda}_t}{\alpha_t} U_{C_0 t}\, MC_t^{HC} \cdot (F_{Kt+1} + 1 - \delta_K) - \frac{\tilde{\lambda}_{t+1}}{\alpha_{t+1}} U_{C_0 t+1}\, MC_{t+1}^{HC} \cdot G_{t+1}. \tag{28}$$

With this definition (27) can be written as

$$\Delta_t = \frac{f + F_{L_0 t}}{G_t^{'}} (F_{Kt+1} + 1 - \delta_K) - \frac{f + F_{L_0 t+1}}{G_{t+1}^{'}} G_{t+1}$$

$$- F_{L_1 t+1} L_{1t} - [F_{L_0 t+1} \cdot (L_{0t+1} - E_{t+1}) - f E_{t+1}]. \tag{29}$$

Comparison of (29) and (8) reveals that Δ_t is the efficient wedge between the social cost and the social benefit of investing in human capital in period *t* instead of postponing the investment by one period. A positive wedge stands for subsidizing relative to the first best. *A priori* the sign of Δ_t is indeterminate. This is different if (28) is evaluated at a balanced growth path. By definition, balanced growth means that the non-leisure choices $L_{0t} = L_0, L_{1t} = L_1$, and $E_t = E$ are constant in *t* while consumption, output, and both types of capital all grow at the common gross rate $G = G(E)$, so that we have $H_{t-1} = G^t H_{-1}, K_{t-1} = G^t K_{-1}, C_{it} = G^t C_{i0}$. At balanced growth $F_{Kt+1} = F_K$, $G_{t+1} = G$ in *t*. Assuming utility to be homogeneous of degree *d* in consumption, the other variables entering (28) take on the following values:

(i) $U_{C_0 t} = G^{(d-1)t} U_{C_0 0} \equiv G^{(d-1)t} U_{C_0}$.

(ii) $$MC_t^{HC} \underset{(17)}{=} -\frac{L_{1t} U_{L_1 t}}{U_{C_0 t}} \frac{1}{G'(E_t) H_{t-1}} \eta'(E_t) = -\frac{L_1 U_{L_1 0} G^{dt}}{U_{C_0 0} G^{(d-1)t}} \frac{1}{G'(E) H_{-1} G^t} \eta'(E)$$

$$= -\frac{L_1 U_{L_1}}{U_{C_0}} \frac{1}{G' H_{-1}} \eta' = MC_0^{HC} \equiv MC^{HC}.$$

Because U is homogeneous of degree d, W is likewise homogeneous of degree d. As a result, the growth factor G^t cancels out in equation (19): $W_{C_0 t} = -\dfrac{W_{L_0 t}}{F_{L_0 t} H_{t-1}}$. After cancelling out, the only variable carrying an index t in this equation is the Lagrange multiplier $\tilde{\lambda}_t$. Hence

(iii) $\tilde{\lambda}_t = \tilde{\lambda}$, and *a fortiori*

(iv) $\alpha_t = W_{C_0 t} = G^{(d-1)t}\, W_{C_0 0} \equiv G^{(d-1)t}\, W_{C_0}$ and $\dfrac{U_{C_0 t}}{\alpha_t} = \dfrac{G^{(d-1)t} U_{C_0}}{G^{(d-1)t} W_{C_0}} = \dfrac{U_{C_0}}{W_{C_0}}$.

Eventually, setting $R \equiv F_K + 1 - \delta_K$, (28) can be written as

$$\Delta = \tilde{\lambda} \frac{U_{C_0}}{W_{C_0}} \cdot MC^{HC} \cdot (R - G). \tag{30}$$

Interpret $\tilde{\lambda} U_{C_0} / W_{C_0}$ as the social cost associated with the implementability constraint. This factor is positive if the implementability constraint is binding, $\tilde{\lambda} > 0$, which is the case if the non-availability of lump-sum taxes is a binding constraint.[4] 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 and increasing in η'. Finally, $R - G$ is the growth gap, which by (10) must be positive as well. Hence Δ is the product of three positive factors.

Proposition 3: Assume selfish behaviour and U to be homogeneous of degree d in consumption. At balanced growth it is second best to subsidize education relative to the first best if the dynamic cost of education, MC^{HC}, is positive. The strength of positive distortion increases in (i) the dynamic cost of education, (ii) the growth gap, and (iii) the cost resulting from the non-availability of lump-sum taxes.

[4] We abstain from proving in detail that the Lagrange multiplier is positive. Jones et al. (1997, p. 109) do this for a maximization which comes close to the present one. The intuition is the following. Paying generation t some positive lump-sum income would show up on the right-hand side of (16). The Lagrange multiplier must be positive if increasing such a lump-sum income can be shown to have a negative effect on the planner's objective function. The effect is indeed negative, because such a lump-sum transfer must be paid at the expense of government funds, which are generated by distortive taxes. Although the government budget constraint is not modelled explicitly, it has to be respected. This follows from Walras's law. In summary, the non-availability of lump-sum taxes is the reason why $\tilde{\lambda}$ is positive.

This is a remarkable result, for reasons explained before. 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. Not so evident is the result 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 even less obvious. 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, however, is true. The key assumption is the increasing elasticity of the human capital investment function with respect to education. If the elasticity is increasing, the dynamic cost of education is positive. With a positive dynamic cost of education it is second best to subsidize education. This has been shown before, by Braun (2009) and Richter (2009), to hold in static analysis, and it is shown here to extend to the dynamic framework. The need to subsidize increases in the factors listed in Proposition 3. In particular, it increases in the derivative of the human capital investment function's elasticity.

We finally turn to the study of labour taxation. Of particular interest is the efficient taxation of nonqualified labour relative to qualified labour. As the definition of W_t in (18) is structurally asymmetric in L_{0t} and L_{1t}, one may easily conjecture that qualified and nonqualified labour should be taxed differently. To make a clear case for differentiated taxation and to obtain clear-cut results, we focus on balanced growth and we assume

$$U \equiv V(C_0, C_1) - \sum_{i=0}^{1} D(L_i) \tag{31}$$

with some function V which is homogenous of degree d. In this particular case the first-order condition (19) implies the following:

$$W_{L_0} + F_{L_0} H_{-1} W_{C_0} = 0. \iff$$

$$(1+\tilde{\lambda}\, d)[U_{L_0} + F_{L_0} H_{-1} U_{C_0}] = \tilde{\lambda}\, [L_0 D''(L_0) + (1-d) D'(L_0)] \ . \tag{32}$$

Similarly, (20) implies

$$W_{L_1} + F_{L_1} G H_{-1} W_{C_1} = 0 \iff$$

$$(1+\tilde{\lambda}\, d)[U_{L_1} + F_{L_1} GH_{-1}\, U_{C_1}] \;=\; \tilde{\lambda}\,[(1-\eta)L_1 D''(L_1) + (1-d-\eta)D'(L_1)]\ . \tag{33}$$

Denote by $1/\nu_i$, $\nu_i \equiv L_i D_i^{''} / D_i^{'}$, the wage elasticity of labour supply, and define tax rates τ_i by setting $(1-\tau_0)F_{L_0}H_{-1} = -\dfrac{U_{L_0}}{U_{C_0}}$, $(1-\tau_1)F_{L_1}GH_{-1} \equiv -\dfrac{U_{L_1}}{U_{C_1}}$ $\Leftrightarrow$ $(1-\tau_i)F_{L_i} = \omega_i$. Dividing (33) through by (32) gives us

$$\frac{\tau_1/(1-\tau_1)}{\tau_0/(1-\tau_0)} = \frac{\nu_1 + 1 - d - \eta(\nu_1+1)}{\nu_0 + 1 - d}\ . \tag{34}$$

For $\eta = 0$ and *d*=1, (34) is the familiar *inverse elasticity rule*. According to this rule, wage taxes should vary inversely with the wage elasticities of labour supplies. This rule is extended by (34) to allow for endogenous education. The effect of education is to reduce the tax on qualified labour relative to the tax on nonqualified labour. The deviation from the inverse elasticity rule increases in the elasticity of the human capital investment function, η. See Richter (2009), who derives the same rule (34) for the static framework and *d*=1. It has to be noted that the given interpretation of (34) assumes a positive numerator. Such positivity is only ensured if the convexity of *D* (as measured by ν_1) and/or the concavity of *V* (as measured by $1-d$) is sufficiently strong to compensate for the lack of concavity of the human capital accumulation equation (1). This lack of concavity is measured by η, and positivity of the numerator requires η to be less than $(\nu_1 + 1 - d)/(\nu_1 + 1)$.

Proposition 4: Assume selfish behaviour, and *U* to satisfy (31). On a balanced growth path it is then second best to tax labour according to the inverse elasticity rule (34). The effect of endogenous education is to reduce the tax on qualified labour relative to the tax on nonqualified labour.

The reader may be interested in learning how second-best policy translates into specific tax and subsidy rates. The problem is that Ramsey tax analysis only lends itself to an implicit determination of policy rates. This is well known from static analysis, and an explicit determination encounters even more difficulties in dynamic analysis. Just for the sake of illustration, consider the special case in which the utility specification (31) holds. Hence Proposition 1 applies, and saving should remain untaxed. Assume that growth is balanced. Proposition 4 applies, and the inverse elasticity rule (34) must hold. Denote by σ the rate by

which the monetary cost of education should be subsidized in second best, $(1-\sigma)f=\varphi$. The optimal set of rates τ_1, τ_2, and σ must solve three equations simultaneously. These equations are (34), the government budget constraint, and the requirement of subsidizing education efficiently. The last means that the extreme hand-sides of (35) are equal:

$$F_{L_0}L_0+(F_{L_0}+f)E\,[\frac{1}{\eta}-1]+\Delta = R\frac{F_{L_0}+f}{G'}-F_{L_1}L_1 = \frac{R}{G'}[\frac{\tau_0-\tau_1}{1-\tau_1}F_{L_0}+\frac{\sigma-\tau_1}{1-\tau_1}f] \quad (35)$$

The first equality characterizes second-best education policy at balanced growth. It follows from (29). The second equality characterizes the optimal private choice of education. It is derived from (15) and (13) after setting $(1-\tau_i)F_{L_i}=\omega_i$, $(1-\sigma)f=\varphi$ and after making some straightforward substitutions. The condition (35) suggests that the need to subsidize 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, $F_{L_0}L_0+(F_{L_0}+f)E(\frac{1}{\eta}-1)$. Note however that this partial analytical interpretation ignores the fact that the efficient policy rates are jointly determined by (34), (35), 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 the sum of discounted lifetime utilities (3) in $C_{0t}, C_{1t}, L_{0t}, L_{1t}, E_t, H_t$, and K_t (t=0,1, ...). This objective is maximized subject to the human capital accumulation constraint (1) and the dynasty's budget constraint,

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

$$= \sum_{t=0}^{\infty}[\pi_t C_{0t}+\pi_{t+1}C_{1t}+\pi_t\varphi_t E_t H_{t-1}+(\pi_t-R_{t+1}\pi_{t+1})K_t] \qquad (\lambda). \qquad (36)$$

The price and cost variables have the same meaning as before. The first-order conditions are (t=0,1, ...)

$$\beta^t U_{C_0t}=\lambda\pi_t,\ \ \beta^t\, U_{C_1t}=\lambda\pi_{t+1},\ \ \omega_{0t}\ H_{t-1}\,U_{C_0t}=-U_{L_0t},\ \ \omega_{1t}H_t\,U_{C_1t}=-U_{L_1t}, \qquad (37)$$

$$\mu_t G_t^{'} = (\varphi_t + \omega_{0t})\, U_{C_0 t}\,, \quad R_{t+1} = \pi_t / \pi_{t+1}\,, \tag{38}$$

$$\lambda \pi_{t+1} [\, \omega_{1t} L_{1t} + \omega_{0t+1} (\, L_{0t+1} - E_{t+1}) - \varphi_{t+1}\, E_{t+1}\,] = \beta^t\, \mu_t - \beta^{t+1}\, G_{t+1}\, \mu_{t+1}\,. \tag{39}$$

The last condition implies

$$\lambda \sum_{t=0}^{\infty} \pi_{t+1} [\omega_{1t} L_{1t} + \omega_{0t+1}\, (L_{0t+1} - E_{t+1}) - \varphi_{t+1} E_{t+1}] H_t \underset{(39)}{=} \sum_{t=0}^{\infty} [\beta^t \mu_t H_t - \beta^{t+1} \mu_{t+1} H_{t+1}]$$

$$= \mu_0 H_0 \underset{(38)}{=} \frac{\varphi_0 + \omega_{00}}{G_0^{'}} U_{C_0 0} H_0\,. \tag{40}$$

Multiplying the budget constraint (36) through by λ and using (37), (38), and (40) to substitute for $\lambda \pi_t, \lambda \pi_{t+1}, \omega_{0t}, \omega_{1t},$ and R_{t+1} in (36) yields the implementability constraint

$$\sum_{t=0}^{\infty} \beta^t \sum_{i=0}^{1} C_{it} U_{C_i t} = B \qquad (\tilde{\lambda}) \tag{41}$$

with

$$B \equiv \{[\omega_{00} (L_{00} - E_0) - \varphi_0 E_0] H_{-1} + \frac{\varphi_0 + \omega_{00}}{G_0^{'}} H_0\} U_{C_0 0}\ .$$

Similarly, (37) and (38) can be used to substitute for $\lambda \pi_{t+1}, \omega_{0t+1}, \omega_{1t},$ and μ_t in (39), which leaves us with (t=0,1, ...)

$$-L_{1t} U_{L_1 t} - \beta\, [(\, L_{0t+1} - E_{t+1}\,) U_{L_0 t+1} + \varphi_{t+1}\, E_{t+1}\, U_{C_0 t+1}\, H_t\,]$$

$$\underset{(39)}{=} \{\, \mu_t - \beta\ \ G_{t+1}\, \mu_{t+1}\} H_t = \mu_t\, H_t - \beta\, \mu_{t+1}\, H_{t+1}$$

$$\underset{(38)}{=} [\, \varphi_t\, U_{C_0 t} - U_{L_0 t}\, \frac{1}{H_{t-1}}\,] \frac{H_t}{G_t^{'}} - \beta\, [\, \varphi_{t+1}\, U_{C_0 t+1} - U_{L_0 t+1}\, \frac{1}{H_t}\,] \frac{H_{t+1}}{G_{t+1}^{'}}\ . \qquad (\gamma_t \beta^t) \tag{42}$$

The planner maximizes the sum of discounted lifetime utilities (3) in $C_{0t}, C_{1t} L_{0t}, L_{1t}, E_t, H_t, K_t$, and φ_t (t=0,1, ...) subject to the resource constraint (2), the accumulation constraint (1), and the behavioural constraints (41) and (42). It is important to note that the monetary cost of education φ_{t+1} (t=0,1,...) only appears explicitly in the condition (42). By contrast, the planner's objective function and the constraints (1), (2), and (41) are independent of φ_{t+1}. The condition (42) can therefore be treated as a relationship by which the "free" policy variable φ_{t+1} can be determined. This solution procedure is feasible because the coefficient of φ_{t+1} in

(42) does not vanish. The coefficient equals $\beta U_{C_0 t+1}[E_{t+1}H_t - \frac{H_{t+1}}{G'_{t+1}}]$ = $\beta U_{C_0 t+1}E_{t+1}H_t(1-\frac{1}{\eta_{t+1}})$ < 0. 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 (41). The same kind of solution procedure has been applied 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 implementability constraint (41) or which drop out when making particular assumptions. The optimization with respect to those variables is not affected by (41) and should therefore remain undistorted.

Proposition 5: Assume altruistic behaviour. Then it is second best not to distort education for all generations except the first.

Proposition 6: Assume altruistic behaviour, and the utility function to be weakly separable between consumption and non-leisure and homothetic in consumption. Then it is second best not to distort the accumulation of capital for all generations except the first.

Proposition 7: Assume altruistic behaviour, and the utility function to be additive separable between consumption and non-leisure. Then it is second best to tax qualified and nonqualified labour uniformly if, and only if, it is second best not to distort the accumulation of capital. This holds for all generations except the first.

The proof of Proposition 5 is rather straightforward. Just note that the variables E_t, H_t, K_t (t>0) do not enter the implementability constraint. Taking partial derivatives of the Lagrange function with respect to these variables and substituting for the Lagrange multipliers μ_t, α_t yields the efficiency condition (8) for t>0. The proof of Proposition 6 parallels the one of Proposition 1 and is therefore skipped. The proof of Proposition 7 is as follows. Set

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

Additive separability of U between consumption and non-leisure implies $W_{L_i t} = U_{L_i t}$ (i=0,1). Hence the social and private marginal rates of intertemporal substitution in the supply of non-leisure are equal,

$$\frac{W_{L_1 t}}{W_{L_0 t}} = \frac{U_{L_1 t}}{U_{L_0 t}} \underset{(37)}{=} \frac{\omega_{1t} H_t U_{C_1 t}}{\omega_{0t} H_{t-1} U_{C_0 t}}. \tag{43}$$

Taking partial derivatives of the Lagrange function with respect to K_t, L_{0t}, L_{1t}, yields (21) and $W_{L_0 t} = -\alpha_t F_{L_0 t} H_{t-1}$, $W_{L_1 t} = -\alpha_{t+1}\beta F_{L_1 t+1} H_t$ (t>0). Therefore, (43) $\Leftrightarrow$

$$\frac{\alpha_{t+1}\beta F_{L_1 t+1} H_t}{\alpha_t F_{L_0 t} H_{t-1}} = \frac{\omega_{1t} H_t U_{C_1 t}}{\omega_{0t} H_{t-1} U_{C_0 t}} \Leftrightarrow \frac{F_{L_1 t+1}}{F_{L_0 t}} \underset{(21)}{=} [F_{Kt+1} + 1 - \delta_K] \frac{\omega_{1t}}{\omega_{0t}} \frac{U_{C_1 t}}{U_{C_0 t}}.$$

Define tax rates τ_{it} by setting $1-\tau_{1t} \equiv \omega_{1t} / F_{L_1 t+1}$, $1-\tau_{0t} \equiv \omega_{0t} / F_{L_0 t}$. Hence, (43) $\Leftrightarrow$

$$\frac{1-\tau_{0t}}{1-\tau_{1t}} = [F_{Kt+1} + 1 - \delta_K] \frac{U_{C_1 t}}{U_{C_0 t}}, \tag{44}$$

which reveals that taxes on labour are independent of age if the right-hand side equals one. The latter clearly holds if, and only if, it is second best not to distort saving.

Proposition 6 is just what one would expect in view of the literature. 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 life cycle. The intuitive explanation for this result is that labour supplied in the second life period differs from labour supplied in the first period. While Proposition 4 confirms the result of Erosa and Gervais on assuming selfish individuals, Proposition 7 does not. Obviously, in the present framework altruism removes the need to employ age-dependent labour taxes for descendent generations in situations where the planner would not distort saving. Age-dependent labour taxes are then used only as a correcting device if it is second best to distort saving. This is an intriguing result. Above, it is derived from the equality of the social and private marginal rates of intertemporal substitution in the supply of non-leisure, (43). For this equality to hold we have to assume not only altruism, but also a sufficiently rich set of policy instruments. In particular, the planner must be able to choose ω_{it} independently of φ_t. In other words, the planner must be able to optimize the taxation of labour separately from the subsidization of education.

Finally, Proposition 5 is interesting in that it is much stronger than the results derived in the Chamley-Judd literature. It holds for arbitrary utility functions, and it does not assume balanced growth. That is, Proposition 5 is logically stronger than Propositions 6 and 7. And it is also much stronger than Proposition 2, which assumes the human capital investment function to be isoelastic. By contrast, Proposition 5 even holds for functions *G* which fail to be isoelastic. All this strongly 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 if no lump-sum income accrues 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 implementability constraint (41) for *t*>0. Furthermore, the only lump-sum income modelled is income earned by the parent generation living in period 0. On setting $\pi_0 = 1$, this income equals $F_{K0}K_{-1} + F_{L_1 0}L_{1-1}H_{-1}$ + $(1-\delta_K)K_{-1}$. It does not show up in the dynasty's budget constraint (37). It must therefore be income accruing to the government budget. The Production Efficiency Theorem is applicable, and Proposition 5 can be considered to be just a corollary.

The recommendation not to distort education is not easily translated into explicit tax and subsidy rates. The reason is that private incentives are affected by a whole set of tax and subsidy rates, which all must be optimally set. Just inspect the altruist's first-order condition (39) determining the optimal amount of human capital. After substituting for the Lagrange multipliers one obtains

$$\omega_{1t}L_{1t} + \omega_{0t+1}\,L_{0t+1} - (\,\omega_{0t+1} + \varphi_{t+1}\,)\,E_{t+1} \;=\; R_{t+1}\,\frac{\varphi_t + \omega_{0t}}{G_t^{'}} - G_{t+1}\,\frac{\varphi_{t+1} + \omega_{0t+1}}{G_{t+1}^{'}}\,. \tag{45}$$

This condition reveals that the altruist's incentive to invest in human capital is not only affected by taxes on own labour income and the subsidy paid to own direct costs of education. It is additionally affected by the tax on savings, by the next generation's tax on nonqualified labour, and finally by the subsidy paid to the next generation's direct costs of education. More can be said only after making special assumptions. Just for the sake of illustration, assume that utility is homothetic in consumption and additive separable between consumption and non-leisure. Hence Propositions 6 and 7 apply, and it is optimal not to tax saving, $R_{t+1} = F_{Kt+1} + 1 - \delta_K$, and to tax labour independently of age, $1-\tau_t \equiv \omega_{1t} / F_{L_1 t+1} = \omega_{0t} / F_{L_0 t}$ (*t*>0). If optimal wage taxes do not differentiate across generations, $\tau_t = \tau$, then, and only then, can one infer that it is compatible with efficiency for the direct cost of education to be subsidized

at the same rate as labour income is taxed, $\varphi = (1-\tau) f$. This follows immediately from comparing (45) with (8). If the mentioned assumptions do not hold, it is difficult to make definite statements about the efficient structural relationship between labour tax rates and education subsidy rates.

The government has to finance the exogenous cash flow of government expenditures A_t (t>0). 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 implementability constraint (41) is binding, and it cannot be ruled out that it is efficient to distort the choice of education of generation 0. This raises the question of how to design optimal human capital policy for generation 0. 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, generation 0's education should not be distorted if the human capital investment function is isoelastic. If however the dynamic cost of education is positive, education should be positively distorted relative to the first best. To show this we maximize (3) subject to (1), (2), (41), and (42). Taking partial derivatives of the Lagrange function yields the following results after some tedious but straightforward manipulations have been made:

$$\frac{\partial}{\partial \varphi_0}: \quad \gamma_0 = -\tilde{\lambda}(1-\eta_0), \tag{46}$$

$$\frac{\partial}{\partial \varphi_1}: \quad \gamma_1 = \gamma_0 (1-\eta_1), \tag{47}$$

$$\frac{\partial}{\partial E_0}: \quad \frac{\mu_0}{\alpha_0} \underset{(44)}{=} \frac{f + F_{L_0 0}}{G_0^{'}} - \frac{\tilde{\lambda}}{\alpha_0} U_{C_0 0} \frac{\varphi_0 + \omega_{00}}{G_0^{'}} [1 + \frac{E_0 G_0^{''}}{G_0^{'}}], \tag{48}$$

$$\frac{\partial}{\partial E_1}: \quad \frac{\mu_1}{\alpha_1} \underset{(44),(45)}{=} \frac{f + F_{L_0 1}}{G_1^{'}} - \frac{\tilde{\lambda}}{\alpha_1} U_{C_0 1} \frac{\varphi_1 + \omega_{01}}{G_1^{'}} [1 - \frac{E_0 G_0^{'}}{G_0}][1 + \frac{E_1 G_1^{''}}{G_1^{'}}]. \tag{49}$$

The first-order condition with respect to K_0 is the same as (21) for t=0. By making use of (46)–(49) and (21) for t=0 we end up with

$$\frac{\partial}{\partial H_0}: \quad \Delta_0 = \frac{f + F_{L_0 0}}{G_0^{'}} (F_{K1} + 1 - \delta_K) - \frac{f + F_{L_0 1}}{G_1^{'}} G_1 - F_{L_1 1} L_{10} - [F_{L_0 1} \cdot (L_{01} - E_1) - f E_1], \tag{50}$$

where

$$\Delta_0 \equiv \frac{\tilde{\lambda}}{\alpha_0} U_{C_0 0} \cdot MC_0^{HC} \cdot (F_{K1} + 1 - \delta_K) - \frac{\tilde{\lambda}(1-\eta_0)}{\alpha_1} U_{C_0 1} \cdot MC_1^{HC} \cdot G_1 \tag{51}$$

and

$$MC_t^{HC} \equiv \frac{\varphi_t + \omega_{0t}}{G_t^{'}} [1 - \frac{E_t G_t^{'}}{G_t} + \frac{E_t G_t^{''}}{G_t^{'}}] = \frac{\varphi_t + \omega_{0t}}{G_t^{'}} \frac{E_t \eta_t^{'}}{\eta_t} \quad (t=0,1) . \tag{52}$$

The variables Δ_0 and MC_t^{HC} are defined so that the parallels with (28) and (17) show up. As MC_t^{HC} vanishes for isoelastic $G(E_t)$, we obtain

Proposition 8: Assume altruistic behaviour and the human capital investment function G to be isoelastic. Then it is second best not to distort the first generation's educational choice.

Proposition 8 is just the altruistic analogue to Proposition 2. It is a result that one could easily conjecture. Altruism goes beyond selfishness in internalizing efficiency effects. If it is second best not to distort education given that G is isoelastic and behaviour selfish, then it should all the more be second best not to distort education given that G is isoelastic and behaviour altruistic.

Things are less straightforward if the dynamic cost of education is positive. Without making further assumptions, it is difficult to sign Δ_0. However, we are able to derive a direct analogue to Proposition 3. More precisely, Δ_0 can be shown to be positive if the growth path is balanced and if utility is homogeneous of degree d. The assumption of balanced growth has the effect of neutralizing the impact of initialization.

The proof is only sketched. First note that $\omega_{0t} = \omega_0$ follows from (37). In a second step G^{dt} is shown to be a factor that cancels out of the constraint (42), so that φ_t and φ_{t+1} are the only remaining variables in (42) carrying an index t. The equation can then be used to solve for $\varphi_t = \varphi_{t+1} \equiv \varphi$. This is a feasible procedure, as the coefficient of φ does not vanish. Just note that after dividing through by G^{dt} the coefficient equals $\beta G^d U_{C_0 0}[EH - \frac{GH}{G^{'}}] + U_{C_0 0} \frac{GH}{G^{'}}$ = $U_{C_0 0}\, EH[\, \beta G^d + \frac{1}{\eta}(1 - \beta G^d\,)]$. The condition of transversality, $\beta G^d < 1$, implies that the coefficient is positive. Plugging φ into (50) yields $MC_t^{HC} = MC^{HC}$. In order to derive $\Delta_0 = \Delta > 0$, assume $MC^{HC} > 0$ and note

$$\frac{\tilde{\lambda}}{\alpha_0} U_{C_0 0} \cdot (F_{K1} + 1 - \delta_K) > \frac{\tilde{\lambda}(1-\eta)}{\alpha_1} U_{C_0 1} \cdot G \underset{(21)}{\Leftrightarrow} U_{C_0 0} > \beta(1-\eta)\, G^{d-1} U_{C_0 0} \cdot G$$

$$\Leftrightarrow 1 > (1-\eta) \cdot \beta\, G^d .$$

The last inequality follows from η <1 and, once more, from the condition of transversality.

Proposition 9: Assume altruistic behaviour, and U to be homogeneous of degree d in consumption. At balanced growth it is second best to subsidize the first generation's educational choice relative to the first best if the dynamic cost of education, MC^{HC}, is positive.

It would be nice if one could similarly characterize second-best policy with regard to the first generation's choice of labour and saving. However, analogues to Propositions 1 and 4 seem not to hold. In particular, it seems that the first generation's saving decision is systematically distorted. The reason is the factor $U_{C_0 0}$ entering the right-hand side of (41). This factor implies a lack of symmetry when taking partial derivatives of B with respect to C_{i0} (i=0,1). As a result it is second best to distort saving.

The parallelism between Proposition 9 and Proposition 3 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 the second-best tax policy for descendent generations is more like the first-best policy. The accumulation of human capital should remain undistorted, and – if utility functions are well selected – labour taxes need not be differentiated across the individual life cycle. The short-run policy recommendations for altruism, however, parallel the long-run recommendations for selfishness. Labour has to be taxed, and – given that the elasticity of the human capital investment function is increasing – education should be subsidized relative to the first best. Whether saving should be taxed is not a matter of selfishness or altruism. With regard to descendent generations 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 in the present study. Such a procedure is, no doubt, debatable. Critical is the ignoring of possible causes 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 of studying efficient taxation in Ramsey's tradition. More precisely, this paper aims at bridging the gap that separates the two strands of Ramsey tax analyses which exist for the finite and the infinite planning horizon. Our knowledge of efficient human capital policy in Ramsey's tradition is largely shaped by incompatible results derived from static and dynamic analyses. 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 efficient policy does not differentiate between human and nonhuman capital. By way of contrast, static analysis strongly suggests differentiated policies. Whether education should be distorted or not appears to depend primarily on 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 be not distorted by second-best policy (Jacobs and Bovenberg, 2008). By way of contrast, the question of whether saving should be distorted or not primarily has to be answered with regard to the taxpayer's consumption preferences. More precisely, saving should not be taxed if the taxpayer's utility is weakly separable between consumption and labour/non-leisure and homothetic in consumption (Atkinson and Stiglitz, 1972).

The model filling the gap between finite and infinite Ramsey tax analyses is one with overlapping generations. The present paper studies second-best policy for education, saving, and labour in such an overlapping-generations model with 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 by Atkeson, Chari, and Kehoe (1999) on one side and by Wigger (2002, Sec. 3.4) and Docquier et al. (2007) on the other side. 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 work of Wigger and Docquier et al. does integrate them. However, it does not allow for 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 where Wigger and

Docquier et al. stop. It goes beyond Atkeson et al. by integrating endogenous education and overlapping generations, and it goes beyond Wigger and Docquier et al. by endogenizing labour supply and by doing second-best tax analysis.

The present paper studies two possible reasons for allocational inefficiency. One is the non-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 Wigger and by Docquier et al., selfishness is the source of an intergenerational externality. It gives reason to subsidize education relative to *laissez-faire*. Such subsidization, however, requires government revenues. In the framework studied by Wigger and by Docquier et al. it is efficient to subsidize education 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. As it turns out, it is still second best not to distort education if only the investment function is isoelastic. This result can be considered to be the dynamic version of the education efficiency proposition known from static Ramsey analysis.

It is, however, argued that an isoelastic investment function has the unappealing implication that all human capital accumulated by past generations melts down to zero if only one generation stops investing. If, by way of contrast, human capital depreciates just by some fraction and if the investment function's elasticity is increasing, then investment incentives should exceed the first-best level at balanced growth. In other words, it is efficient in the long run to combine positive tax wedges in the labour market with a negative effective tax wedge for education. The need to subsidize is shown to increase in (i) the dynamic cost of education, (ii) the cost resulting from the non-availability of lump-sum taxes, and (iii) the growth gap. Furthermore, it turns out to be efficient to tax labour so such that qualified labour is less distorted than nonqualified labour.

If taxpayers are altruists with respect to descendent generations, one clear reason for government intervention does not apply. 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 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 whenever it is second best not to distort saving. This 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. In the present framework, however, altruism has the effect of implying equality of the social and private marginal rates of intertemporal substitution in the supply of non-leisure. The optimality of uniform labour taxation is an immediate though intriguing corollary to this equality. In view of the Chamley-Judd literature, results suggesting non-distortionary taxation may not be too surprising.

Striking, however, is the strength of the result concerning human capital accumulation. While the other results on non-distortionary taxation require specific utility functions, the result on human capital accumulation holds without any comparable qualification. One only has to assume that no lump-sum income 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 every distortion. 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. Nor do the results on non-distortionary taxation extend to the dynasty's first generation, indexed by zero in the present paper. A more precise characterization of optimal policy for generation 0 is difficult, as the specific features not only depend on the shape of the human capital investment function but also on initial values of key variables. As in the case with selfish individuals, it is efficient not to distort education if the investment function is isoelastic in education. If, however, the elasticity is increasing and if the impact of initialization is suppressed by assuming balanced growth, it is second best to subsidize education relative to the first best. The reason is the same as the one given before in the scenario with selfish individuals. An increasing elasticity of the investment function has the effect that it is second best to subsidize education in static analysis, and this effect extends to the dynamic framework. At balanced growth the need to subsidize increases in the derivative of the investment function's elasticity, and it is the stronger, the more binding the non-availability of lump-sum taxes is and the more deficient growth is.

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 – given that the elasticity of the human capital investment function is increasing – education should be subsidized 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 intertemporal 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.

Braun, C., 2009, Optimal taxation of education with an initial endowment of human capital, TU Dortmund University, mimeo.

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.

Collins K. A. and J. B. Davies, 2004, Measuring effective tax rates on human capital: methodology and an application to Canada, in: Measuring the tax burden on capital and labor, P. B. Sörenson, ed., MIT Press, 171–211.

Del Rey, E. and M. del Mar Racionero, 2002, Optimal educational choice and redistribution when parental education matters. Oxford Economic Papers 54, 435–448.

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, 2007, Optimal accumulation in an endogenous growth setting with human capital. Journal of Economic Theory 134, 361–378.

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.

Lucas, R. E., 1988, On the mechanics of economic development. Journal of Monetary Economics 22, 3–42.

OECD (2009 and 2005), Education at a Glance.

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., 2009, Taxing education in Ramsey's tradition. Forthcoming in: Journal of Public Economics.

Sandmo, A., 1974, A note on the structure of optimal taxation. American Economic Review 64, 701–706.

Trostel, P. A., 1996, Should learning be subsidized? Public Finance Quarterly 24, 3–24.

Uzawa, H., 1965, Optimum technical change in an aggregate model of economic growth. International Economic Review 6, 18–31.

Wigger, B.U., 2002, Public pensions and economic growth. Springer-Verlag.