Efficient Tax Policy Ranks Education Higher Than Saving[+]

by

Wolfram F. Richter

University of Dortmund, CESifo Munich, and IZA Bonn

August, 2007

This version: June, 2008

Abstract:

Assuming a two-period model with endogenous choices of labour, education, and saving, it is shown to be second-best efficient to distort qualified labour less than nonqualified labour and, most notably, not to distort the choice of education. In general, the latter implies distorting the saving decision. Hence a strict order of policy priority is derived. Efficient tax policy ranks education higher than saving. All the results hold for arbitrary utility functions and only the one on efficient education assumes an isoelastic earnings function. Assuming isoelasticity is however justified with reference to the empirically well-founded Power Law of Learning.

Address:

Wolfram F. Richter, University of Dortmund, Department of Economics, 44221 Dortmund, Germany; email: Wolfram.Richter@tu-dortmund.de, phone: +49-231-755-3146, fax: +49-231-755-5404

---

1. Introduction

One of the most debated issues in public economics is whether capital income should be taxed or not. Although there seems to be broad consensus that taxing savings is presumably "a bad idea" (Atkeson et al., 1999), one has to admit that truly compelling arguments are lacking. While equity arguments are used both to attack and to justify the taxation of capital, efficiency arguments are at best inconclusive. They lend themselves to argument against the taxation of capital in a model of finite horizon only if particular preferences are assumed for which the empirical evidence is weak. If the choice of preferences is to play no critical role, one has to rely on Chamley (1986) and Judd (1985). Their result establishes that the optimal tax rate on capital income tends to zero in a model of infinite horizon. The problem with this result, however, is not only that it does not extend to a finite horizon. More critical is that it is biased in one important respect. It holds for a framework that models the accumulation of nonhuman capital but ignores the accumulation of human capital. If however the two courses of capital accumulation are modelled more symmetrically, the reason for discriminatory taxation disappears. More precisely, Chamley and Judd's result on zero capital-income taxation in the limit is then seen to extend to labour taxes (Jones et al., 1997). This result is not convincing either, however. It relies on blurring differences between human and nonhuman capital, and it raises even more the question as to which ultimate economic feature makes the one differ from the other.

In this paper, the difference between human and nonhuman capital is modelled as a short-run difference in the returns to individual acquirement. By assumption, nonhuman capital is acquired by saving at constant returns, while human capital is acquired by education (learning) at decreasing returns. Decreasing returns give rise to pure profit. The profit of education is special for two reasons. First, the immediate outcome of education is ability only. Skills acquired by education do not pay off if not combined with additional own labour effort. It is in this sense that the profit of education is quasi-pure. Secondly, it is not clear what efficient taxation should do with the quasi-pure profit of education. Consumption and labour taxes cut off some profit; they do not however skim it off fully. The present paper investigates the efficient

taxation of education and it does so within the standard two-period life-cycle model with endogenous choices of education, consumption, labour, and saving. Two major results are derived. One is Proposition 3 which states that it is second-best not to distort the choice of education but only the other household decisions including saving. This result is interpreted as establishing a strict order of policy priority: Education incentives should be ranked higher than saving incentives and all other supply-and-demand decisions of private households. From a theoretical point of view, the result calls for qualifying the General Theory of Second-Best which has been developed by Lipsey and Lancaster (1956/57) and which seems to suggest spreading distortions uniformly across all feasible margins. The second major result is Proposition 2 which states that qualified labour should be less distorted than nonqualified labour. This result is not only a challenge to any policy aimed at trading off efficiency and equity. It also calls for qualifying Ramsey's (1927) result on second-best taxation which requires equiproportionate compensated reductions in supply-and-demand decisions of private households. All results are shown to hold for arbitrary utility functions. It is only Proposition 3 which assumes the earnings function to be isoelastic.

The paper is structured as follows. Section 2 sets up the model of a representative taxpayer. Section 3 introduces policy instruments and the planner's objective function. Section 4 provides the results on second-best taxation. Section 5 explores the implication for the taxation of saving. Section 6 summarizes. Major proofs are relegated to a technical Appendix.

2. A representative-household model

Consider a representative household living for two periods. Lifetime utility is given by $U(C_1, C_2, L_1, L_2)$, where C_i is consumption and L_i is non-leisure time in period $i=1,2$. Non-leisure time L_2 is identical with second-period labour supply. By contrast, only $L_1 - E$ is time spent in the market; time E is spent on education. First-period labour supply earns a constant wage rate ω_1; the productivity of second-period labour depends on the amount of education. It is paid $\omega_2 H(E)$, where ω_2 is constant while

the *earnings function H(E)* displays positive but diminishing returns, $H'>0>H''$. We will also refer to *qualified labour* in the case of L_2 and of *nonqualified non-leisure* in the case of L_1. Education has a cost in *forgone* (*wage*) *earnings*, which is captured by $\omega_1 E$. *Monetary costs of education* like college fees come on top of these and are modelled by φE. The share of first-period income that is neither spent on education nor on consumption is saved,

$$S = \omega_1(L_1 - E) - \varphi E - qC_1 = \omega_1 L_1 - (\omega_1 + \varphi)E - qC_1. \tag{1}$$

Second-period consumption is constrained by income earned,

$$qC_2 = \rho S + \omega_2 H(E)L_2 . \tag{2}$$

q is the consumer price of consumption and ρ is the gross rate of return to saving. All prices are after taxes and subsidies, and the question is which combination of taxes and subsidies is second-best efficient. The representative household is assumed to maximize utility in C_1, C_2, L_1, L_2, E subject to the lifetime budget constraint

$$\rho qC_1 + qC_2 = \rho\omega_1 L_1 + \omega_2 H(E)L_2 - \pi E \tag{3}$$

stated in second-period units. Interpret $\pi \equiv \rho(\omega_1 + \varphi)$ as the *effective* (*unit*) *cost of education*. As stressed by Heckman (1976), the cost is seen to decrease in ρ.

The analysis relies on the dual approach to optimal taxation. This means that the focus is shifted from the household's (indirect) utility function to its (net) expenditure function. The task of minimizing (net) expenditures subject to an exogenous utility constraint is best solved in a two-step approach. At the first step income derived from education is maximized while keeping the level of L_2 fixed. Let this income be denoted by $Y(\omega_2, \pi, L_2) \equiv \max_E[\omega_2 H(E)L_2 - \pi E]$, and the optimal amount of education by $E(\omega_2, \pi, L_2)$. The optimal amount is implicitly defined by the first-order condition, $\omega_2 H'L_2 = \pi$. If the second-period labour supply L_2 were exogenous, Y would stand for pure profit. However, the focus is here on an endogenous choice of L_2. Hence Y has to be interpreted as quasi-pure profit, the source of which is education and its diminishing return. Note that Y is a monotone increasing, convex function of L_2:

$$\frac{dY}{dL_2} = \frac{\partial Y}{\partial L_2} = \omega_2 H(E) > 0, \quad \frac{d^2Y}{dL_2^2} = \omega_2 H' \frac{dE}{dL_2} = -\omega_2 \frac{H'^2}{H''L_2} > 0.$$

Let the second-period wage rate before taxes be denoted by w_2, and the effective social cost of education (i.e., the effective cost before taxes and subsidies) by $p = r(w_1 + f)$. r is the gross rate of return to saving before taxes and subsidies. Equally f is the monetary unit cost of education before taxes and subsidies. *Education is efficient* if the tax wedge between the marginal social return and the effective social cost,

$$\text{tax wedge} \equiv w_2 H'(E)L_2 - p = \pi\left[\frac{w_2 H' L_2}{\pi} - \frac{p}{\pi}\right] = \pi\left[\frac{w_2}{\omega_2} - \frac{p}{\pi}\right],$$

vanishes. Obviously, the tax wedge vanishes if, and only if, the rates of return before and after taxes and subsidies are equal,

$$\frac{w_2}{p} = \frac{\omega_2}{\pi}. \tag{4}$$

Below, it is shown to be efficient (in the second-best sense) not to compromise on efficiency in education. This may be considered to be an important qualification to the General Theory of Second Best as developed by Lipsey and Lancaster (1956/57). The general theory seems to suggest that efficiency is enhanced if small distortions are spread over learning and the supply of labour. In contrast, Proposition 3 states that efficiency requires distorting all household choices except learning.

The expenditure function is defined as

$$e(q, \omega_1, \omega_2, \rho, \varphi; u) \equiv \min[\rho q C_1 + q C_2 - \rho \omega_1 L_1 - Y(\omega_2, \rho(\omega_1 + \varphi), L_2)]$$

in C_1, C_2, L_1, L_2 such that $U(C_1, C_2, L_1, L_2) \geq u$.

Hotelling's lemma yields $e_q = \rho C_1 + C_2$, where $C_i = C_i(q, \omega_1, \omega_2, \rho, \varphi; u)$ solves the optimization and where the subscript q denotes a partial derivative. By relying on a straightforward generalization of the textbook version of Hotelling's lemma one likewise derives the identities $e_1 \equiv \frac{\partial e}{\partial \omega_1} = -\rho(L_1 - E)$, $e_\varphi = \rho E$, and $e_\rho =$

$qC_1 - \omega_1 L_1 + (\omega_1 + \varphi)E = -S$. Just like C_i, the functions L_i and S are Hicksian ones and have to be evaluated at $q, \omega_1, \omega_2, \rho, \varphi$, and u. As a result, the fully specified education function reads $E = E(\omega_2, \rho(\omega_1 + \varphi), L_2(q, \omega_1, \omega_2, \rho, \varphi; u))$.

3. Policy instruments

The analysis studies the efficient mix of four policy instruments, each of which is distorting. The first one is a tax $t_C > 0$ on consumption. Treating consumption as a numéraire good with a producer price of one, this implies $q = 1 + t_C$.

The second instrument is a tax t_E on the monetary social cost of education f, so that $\varphi = (1 + t_E)f$ results. Negativity of t_E is not ruled out, so that subsidizing education is a feasible policy.

The third instrument is a tax τ on capital income. Again negative values of τ are not ruled out. As it turns out, it is convenient to define the capital tax in *exclusive* form. Denoting the gross social return to saving by r, exclusiveness means that τ satisfies the condition $r = (1 + \tau)\rho$. In other words, the base of the capital tax excludes the tax payment.

The fourth policy instrument is a subsidy σ to labour income earned in the first period. This requires setting $\omega_1 = (1 + \sigma)w_1$, where w_1 denotes the market wage rate. Second-period labour income is assumed to remain untaxed, $\omega_2 = w_2$. Given that consumption is taxable, nothing is gained by introducing a tax on second-period labour. It would only provide a redundant instrument, which could be duplicated by an appropriate choice of the four other policy instruments. As first-period labour is nonqualified while second-period labour is qualified, a positive σ can be interpreted as a policy regime in which labour income is taxed progressively with respect to qualification. All social costs, w_1, w_2, r, and f, are treated as exogenous parameters of the planner's optimization. This holds *a fortiori* for the effective social cost of education, $p = r(w_1 + f)$. Not to endogenize factor prices r and w_2 is only for expositional ease. It can be shown that all results provided in this paper carry over to

the full equilibrium framework if only returns to scale are constant with respect to nonhuman capital and effective labour, $H(E)L_2$.

There is a need to raise tax revenue in order to finance government expenditures. Government's net tax revenue amounts to

$$T \equiv t_C(rC_1 + C_2) + t_E rfE + \tau\rho S - \sigma rw_1(L_1 - E)$$
$$= (q-1)(rC_1 + C_2) + r(\varphi - f)E + (r - \rho)S + r(w_1 - \omega_1)(L_1 - E).$$

By invoking Hotelling's lemma this can be written as

$$T = (q-1)e_q + \frac{1}{\rho}[\frac{q-1}{q}(\rho - r)\omega_1 + r(\omega_1 - w_1)]e_1 + \frac{\rho - r}{q}e_\rho$$

$$+ \frac{1}{\rho}[\frac{q-1}{q}(\rho - r)\varphi + r(\varphi - f)]e_\varphi. \tag{5}$$

In order to derive (5) the various tax bases are transformed and expressed in terms of partial derivatives of the e. The weights assigned to the partial derivatives such as q-1, $\frac{\rho - r}{q}$, and so forth therefore have to be interpreted as transformed tax wedges. The social planner is assumed to maximize tax revenue T subject to the condition that private net expenditure remains constant at zero level, e=0. Let all the conditions of regularity hold that are needed to make the optimization a well-behaved problem and to sustain interior solutions, $C_1, C_2, L_1 - E, L_2, E$ >0. Note that just assuming a quasi-concave utility function would not guarantee interior solutions. Instead, the disutility of qualified labour must be sufficiently convex in order to outweigh the convexity of Y when expressed as a function of L_2. This is borne out in the example $U(C_1, C_2, L_1, L_2)$ = $F(C_1, C_2) - L_1^v - L_2^v$, $H(E) = E^\eta$, η <1 with F being some linear homogeneous function. As shown in Richter (2007) v must exceed $1/(1-\eta)$ if the planner's maximization problem is to be well-behaved. A set of instruments t_C >0, t_E, τ, and σ is said to be second-best efficient if it solves the planner's problem. The combination with t_C >0=t_E=τ=σ is a feasible one. However, it distorts consumption and labour choices, and the key question is whether it is more efficient to reduce t_C and to compensate the effect on utility by appropriate use of the remaining instruments.

4. Efficient policy

Maximizing government's net revenue T in t_C, t_E, τ, and σ subject to a balanced-budget constraint on the taxpayer, $e=0$, yields a problem that can easily be solved by applying the standard Lagrange technique. Maximizing in t_C, t_E, τ, and σ is obviously equivalent to maximizing in q, φ, ρ, and ω_1. In the Appendix it is shown: Taking partial derivatives with respect to the latter variables, invoking Hotelling's lemma, and eliminating the Lagrange multiplier yields a system of three first-order conditions:

$$\frac{\Delta E}{E} = \frac{\Delta L_1 - \Delta E}{L_1 - E} = \frac{\omega_1 \Delta L_1 - q \Delta C_1}{\omega_1 L_1 - q C_1} = \frac{\rho \Delta C_1 + \Delta C_2}{\rho C_1 + C_2}, \tag{6}$$

where the differentiation operator Δ is defined on arbitrary functions $X = X(q, \omega_1, \omega_2, \rho, \varphi; u)$ as follows:

$$\Delta X \equiv (q-1) X_q + \frac{1}{\rho} [\frac{q-1}{q}(\rho - r)\omega_1 + r(\omega_1 - w_1)] X_1$$

$$+ \frac{\rho - r}{q} X_\rho + \frac{1}{\rho} [\frac{q-1}{q}(\rho - r)\varphi + r(\varphi - f)] X_\varphi. \tag{7}$$

According to (7) ΔX is defined as the weighted sum of the partial derivatives of X with the weights given by the transformed tax wedges. It is an approximation of the total change in X when taxes are efficiently chosen. When reshuffling (6) the following set of equalities is obtained:

$$\frac{\Delta E}{E} = \frac{\Delta L_1}{L_1} = \frac{\Delta C_1}{C_1} = \frac{\Delta C_2}{C_2}. \tag{6'}$$

As differentiation is additive, (6') could equally and equivalently be stated in the form where the ratio $\Delta L_1 / L_1$ is replaced with the ratio $\Delta(L_1 - E)/(L_1 - E)$. The system (6') is readily interpreted in the spirit of Ramsey (1927): Efficient policy induces equiproportionate reductions in compensated demands and supplies.

Proposition 1: Efficient policy requires equiproportionate reductions in education, periodic consumption and nonqualified non-leisure/labour.

The striking feature about Proposition 1 is that qualified labour L_2 is not among the listed variables. Indeed, efficiency requires reducing qualified labour less than nonqualified labour.

Proposition 2: Efficient policy requires

$$\frac{\Delta L_2}{L_2} = (1-\eta)\frac{\Delta L_1}{L_1} \quad . \tag{8}$$

In (8), $\eta = EH'/H$ is the elasticity of the earnings function which at this point may well fail to be constant. The proof is relegated to the Appendix. The interpretation of (8) is straightforward. Efficiency requires reducing qualified labour relatively less than nonqualified labour. The ratio equals $1-\eta$ and it decreases in η. In other words, the more elastic the individual earnings function is, the less should qualified labour be reduced in relative terms. Although this makes good sense one must see that it fails to accord with Ramsey's rule of reducing *all* household choice variables equiproportionately. Equiproportionate reductions of qualified and nonqualified labour are efficient only if $\eta=0$. Inelastic earnings have however been precluded by assumption as they are not compatible with an endogenous choice of education.

Propositions 1 and 2 raise the question as to which choices of q, φ, ρ, and ω_1 are second-best. Clearly, one should not expect any interesting relationship to hold in full generality. Still, a remarkably strong result is obtained if the individual earnings function is isoelastic in education, $H(E) = hE^{\eta}$ with $0 < \eta < 1$.

Proposition 3*:* If the individual earnings function is isoelastic, it is efficient not to distort the choice of education.

Given that $\omega_2 = w_2$ has been assumed, not to distort education requires leaving the effective cost unchanged:

$$\rho(\omega_1 + \varphi) = \pi \;=\; p = r(w_1 + f) \,. \tag{4$'$}$$

The proof is given in the Appendix. Proposition 3 has been derived before by Richter (2007) in a less structured model. The generality of the result is striking. Beyond regularity assumptions of household optimization, there are no additional ones needed to constrain the choice of utility functions. However, isoelasticity of the individual earnings function is indispensable.

From a purely mathematical point of view isoelasticity may look very special. However, its assumption can well be defended by referring to the *Power Law of Learning* known from cognitive psychology. The content of the Power Law is the following. According to common experience, most tasks get faster with practice, and this holds across task size and task type. If the relationship between practice and the completion time of a task is plotted, a power law is generally seen to provide the best fit. Education is undoubtedly broader and more complex than the training for certain tasks. However, "the power law of practice is ubiquitous" (Ritter et al., 2001), and it would not be plausible to doubt its empirical relevance for the formation of abilities, which after all is the economically relevant essence of education.

Combining Propositions 1 and 3 implies that efficient policy well tolerates a reduction in educational effort. This reduction cannot be interpreted, however, as a distortion of education, but only as a distortion in the supply of qualified labour. This observation allows one to qualify Trostel (1993) who stresses the negative effect of proportional income and consumption taxation on education. Increasing the consumption tax rate t_C clearly has a negative effect on the choice of E but it does not impact the efficiency condition (4).

The intuition underlying Proposition 3 is the following. In deciding on education and labour supply the private household trades off costs against benefits. The benefits are

given by labour income and by the private profit from education, which in real terms amounts to Y/q. In the general case of $\omega_2 \neq w_2$, the social profit is

$$w_2 H L_2 - pE = w_2 \frac{Y + \pi E}{\omega_2} - pE = \frac{w_2}{\omega_2} Y + w_2 [\frac{\pi}{\omega_2} - \frac{p}{w_2}]E,$$

which equals $w_2 Y / \omega_2$ whenever (4) holds. The significance of efficiency in education is hence seen to be in the alignment of private and social objectives. By maximizing the private profit from education, Y/q, the social profit, $w_2 Y / \omega_2$, is maximized as well. Isoelasticity of H serves to ensure that this perfect alignment of private and social objectives need not be compromised. If the earnings function fails to be isoelastic, it may well be optimal to exploit variations in the elasticity of the earnings function and to trade off resulting efficiency gains against distortions in the choice of education.

Proposition 3 has to be contrasted with a result by Bovenberg and Jacobs (2005, 2001). These authors likewise identify circumstances under which human-capital investment should remain undistorted. See also Jacobs and Bovenberg (2007). There are notable differences, however. Bovenberg and Jacobs focus on the optimal trade-off between equity and efficiency when skill formation is endogenous, and they enlarge the set of policy instruments by assuming that a poll tax is available. In substituting the equity objective for the objective of generating tax revenue, their analysis goes beyond the present one. On the other hand, these authors are only able to derive efficiency of education for a scenario in which the cost of education is purely monetary. Costs of forgone wage earnings are ruled out. It is as if nonqualified non-leisure L_1 were exogenously fixed. As a result, education E cannot be interpreted as learning and isoelasticity of the earnings function cannot be justified with reference to the Power Law of Learning. The versions of efficiency results derived by Bovenberg and Jacobs in various papers require that education E is interpreted as an intermediate production good. Hence efficiency in education has to be regarded as production efficiency and results have to be related to the Production Efficiency Theorem of Diamond and Mirrlees (1971). Other than here, isoelasticity of the earnings function is not needed to prove efficiency in education for the representative taxpayer model. The availability of a poll tax even allows the first best to be sustained in this special case. Isoelasticity is

only needed to sustain second-best solutions if there are heterogeneous taxpayers. By contrast, Proposition 3 has to be interpreted as a proposition on consumption efficiency. It is more a qualification to Lipsey and Lancaster (1956/57) than an extension of Diamond and Mirrlees (1971). It does not assume the availability of a poll tax, and it holds even though education provides disutility and fails to be an intermediate production good. The suggested policy implications are very different, as will become better noticeable when discussing the efficient taxation of entrepreneurial profits. If the source of such profits is a fixed factor, the theorem of Diamond and Mirrlees can only be applied if these profits are taxed at one hundred percent. If however profits result from decreasing returns to learning and require personal labour effort, then Proposition 3 suggests letting a positive share accrue to the entrepreneur.

5. Implications for taxing nonhuman capital

On inserting tax variables, the equality $\rho(\omega_1 + \varphi) = \pi = p = r(w_1 + f)$ is seen to imply

$$\tau = \frac{\sigma w_1 + t_E f}{w_1 + f}. \tag{9}$$

This means that in the second-best optimum the tax on capital income, τ, has to equal the weighted sum of the subsidy to nonqualified labour, σ, and the tax on the monetary cost of education, t_E, with the weights given by the social costs w_1 and f. This tight relationship indicates that one should not expect τ to equal zero except for very special cases. Remember that any efficient set of policy instruments has to solve both (6) and (4′). As a consequence, a pure consumption-tax regime, $t_C > 0 = t_E = \tau = \sigma$, is well feasible and even partially efficient in the sense of not distorting education. As the regime implies $\Delta X = (q-1)X_q$, it is totally efficient, however, only if the elasticities of C_1, C_2, L_1, and E with respect to q happen to be equal at these particular parameter values. In other words, total efficiency requires a well-balanced use of policy instruments, and the efficient τ will only be zero in non-generic cases to be discussed below.

Corollary: It is generically efficient to distort saving.

From combining Proposition 3 and the Corollary an order of policy priority is derived. According to this policy order the decision to acquire human capital ranks higher than the decision to acquire nonhuman capital. The reason is that the former generates quasi-pure profit and the latter does not. Efficiency requires not impairing the incentives to earn ability profit beyond what appears to be unavoidable. Unavoidable are the negative incentives that consumption or wage taxes exert on the supply of qualified labour. By contrast, education should remain undistorted. In general this objective requires distorting the savings decision.

The Corollary has to be contrasted with a result of Jacobs and Bovenberg (2005). These authors demonstrate that there are constellations where a positive tax on nonhuman capital serves to alleviate tax distortions in human-capital investment. The specific feature of this result is that the tax on saving is positive if, and only if, education is distorted. By contrast, the Corollary states that it is generically efficient to distort saving even if education remains undistorted. The result of Jacobs and Bovenberg (2005) is however additional evidence of the fundamental proposition that efficient tax policy ranks education higher than saving. This is so as the tax on saving is justified with regard to a tax distortion in human-capital investment. Such a result however reverses the intuition conveyed by Nielsen and Sörensen (1997). These authors take the capital income tax as given, and they study its implication for efficient taxation of human-capital investment. Hence they implicitly reverse the suggested policy order. Human-capital policy is used to accommodate distortions in nonhuman-capital policy. More precisely, Nielsen and Sörensen show that a progressive income tax is needed to offset the tendency of a proportional comprehensive income tax to discriminate in favour of human-capital investment. In view of Proposition 3, their result is at best one characterizing third-best policy.

One should mention that Nielsen and Sörensen (1997) ignore monetary costs of education when modelling endogenous labour supply, while Jacobs and Bovenberg (2005) ignore forgone wage earnings. Proposition 3 cannot, however, be expected to hold in full generality if just one of these two kinds of education costs is modelled.

The set of prices and policy instruments must be sufficiently rich to neutralize the effect that taxation has on human-capital investment. More precisely, it must be possible to tax or to subsidize the monetary cost of education, and it must also be possible to tax or to subsidize nonqualified labour at a rate that may deviate from both the former and the tax applied to qualified labour.

The Corollary does not rule out the possibility that it is efficient to set $\tau = 0$ for special utility functions. As a matter of fact, efficiency of $\tau = 0$ holds if the utility function takes the form $U(C_1, C_2, L_1, L_2) = U(F(C_1, C_2), L_1, L_2)$ with some homothetic function F. The argument is as follows. If F is homothetic, the optimizing individual allocates lifetime consumption in such a way that the ratio C_1 / C_2 only depends on the rate of interest. Hence $\rho C_1 + C_2 = c(\rho) C_2$ holds with some appropriately specified function $c(\rho)$. On normalizing money units ($r=1$) and setting $\rho \equiv r$, $c(\rho) C_2 \equiv C$, the taxpayer's problem reads

$$\max \tilde{U}(C, L_1, L_2) \quad \text{in } C, L_1, L_2, E$$

$$\text{subject to} \quad qC = \omega_1 L_1 + \omega_2 H(E) L_2 - (\omega_1 + \varphi) E. \tag{3'}$$

Just as before, the social planner maximizes tax revenue. The only change is that the capital tax is no feasible policy instrument. When maximizing revenue with respect to $t_C > 0$, t_E, σ, and $\omega_2 = w_2$, it likewise turns out to be efficient not to distort education. For the details see Richter (2007). In other words, Proposition 3 continues to hold. The intuition is that the policy instrument τ is redundant when the subutility of consumption, $F(C_1, C_2)$, is homothetic. The tax on capital can be set equal to zero without jeopardizing efficiency. In this particular case it is even possible to sign t_E and σ. More precisely, as shown by Richter (2007), it is efficient to levy a *positive* tax on the monetary cost of education and to tax labour *regressively* with respect to qualification. That is, efficiency requires $t_E > 0 = \tau > \sigma$ when $t_C > 0$, $\omega_2 = w_2$. The only additional assumption needed is that the elasticity of consumption exceeds the elasticity of nonqualified non-leisure when q is varied. In other words, the direct effect that a change in q has on consumption needs to be stronger than the indirect effect that such a change has on the supply of nonqualified non-leisure. The intuition for

$t_E > 0 > \sigma$ is that it is efficient to set incentives so that qualified labour is substituted for nonqualified labour. Given that $\omega_2 = w_2$ is assumed to hold, this objective is reached by setting $\omega_1 < w_1 \Leftrightarrow \sigma < 0$. To tax nonqualified labour, however, means to reduce the cost of forgone earnings, the partial effect of which is to encourage human-capital investment. The efficiency condition $\omega_1 + \varphi = w_1 + f$ can then be restored only if the monetary cost of education is positively taxed: $\varphi > f \Leftrightarrow t_E > 0$.

It is an interesting, though still open, question to what extent $t_E > 0 > \sigma$ continues to hold in the non-homothetic case of $U(C_1, C_2, L_1, L_2)$. Equally, it is not clear under which fully general conditions a subsidy to capital, $\tau < 0$, can be shown not to be efficient.

6. Summary

One of the most debated issues in public economics is whether capital income should be taxed or not. This paper suggests that proposed answers may well be useless if human capital, its acquirement, and its taxation are not appropriately taken into account. If this is done, a strict order of policy priority can be derived. According to this order learning ranks higher than saving. The former should not be distorted, which in general requires that the latter will be distorted. The reason for this asymmetric treatment is that education – other than saving – generates quasi-pure profit, and taxation should not impair the efficient generation. This result is remarkably robust. In this paper it has been proved for the standard two-period life-cycle model of a representative household with endogenous consumption, labour, and education. The result does not assume particularly selected utility functions, but only an isoelastic earnings function and a sufficiently rich set of policy instruments. Isoelasticity of earnings is justified with reference to the empirically well-founded Power Law of Learning.

The reader may be inclined to question the asymmetric treatment of education and saving driving the results. It makes good sense, however, to assume decreasing returns to learning but constant returns to saving. Learning is a personal activity which cannot

be delegated to other individuals. Its product has to be combined with own labour. This is quite different with the investment of savings. This activity can be delegated to other individuals ("investment bankers") who have learnt to do the job efficiently and to earn the market rate of return. Allowing for general equilibrium effects induced by decreasing factor returns at the aggregate level does not change the efficient tax structure. It can be shown instead that all results derived in this paper perfectly extend to the closed-economy scenario where the market rate of return to saving, r, and the wage rate earned by qualified labour, w_2, are endogenized. The only assumption needed is constant returns to scale with respect to non-human capital and effective qualified labour.

The present paper is not the first attempt to study economic implications of decreasing returns to learning. Well known is Arrow's (1962) attempt to develop a theory of technical change and growth by drawing on the Learning Curve. In his model, however, the Learning Curve takes the role of a labour demand curve. Knowledge is completely embodied in capital, and at each moment of time capital goods of different vintages are in use. As Arrow stresses himself in his closing comments, the implicit assumption is that learning takes place only as a by-product of ordinary production. By way of contrast, learning is central in the present model. It is an individual investment in one's own productivity and the result of endogenous choice.

One cannot summarize without qualifying the obtained results. There are two points of weakness that need to be addressed more than others. One obvious shortcoming of the present analysis is certainly its pure focus on efficiency. Equity considerations have been entirely ignored. This is different from much of the cited literature and from the work of Bovenberg and Jacobs in particular. The conflict between equity and efficiency in human-capital policy is however not a simple one. In any case, it cannot be discussed satisfactorily just in passing. The interested reader is asked instead to refer to the literature as surveyed, e.g., by Carneiro and Heckman (2003).

The other shortcoming can be seen in the lack of balance in the derived efficiency result. The clear policy prescription not to distort the educational choice contrasts with the inability to derive equally clear policy prescriptions for other areas of taxpayers' choice. Although Proposition 2 definitely calls for reducing qualified labour less than

non-qualified labour it is generally unclear what this means in terms of tax rates. Only if the utility of consumption is assumed to be homothetic and separable from leisure can more be said. In this case, efficient policy relies on taxing nonqualified labour more heavily than qualified labour and on restoring efficiency in education by taxing its monetary cost. See Richter (2007), where the result is interpreted by referring to the (weak) double-dividend hypothesis known from environmental economics. No doubt, popular conceptions of good education policy look different (Trostel, 2002). The contrast is however nothing but another indication for the need to enrich the present analysis by equity considerations in future research.

7. Appendix

The *proof of* (6) relies on taking partial derivatives of the Lagrange function $T - \lambda e$ with respect to q, φ, ρ, and ω_1:

$$\frac{\partial}{\partial \varphi}[T - \lambda e] = 0 \quad \Leftrightarrow \quad (\lambda - 1)e_\varphi + \frac{\rho - r}{q\rho}e_\varphi = \Delta e_\varphi . \tag{10}$$

By Hotelling's lemma and by the definition of the Δ-operator, one obtains

$$e_\varphi = \rho E \quad \text{and} \quad \Delta e_\varphi = \Delta(\rho E) = \rho \Delta E + \frac{\rho - r}{q}E . \tag{11}$$

Plugging (11) into (10) yields $\lambda - 1 = \Delta E / E$. Similarly one derives

$$\lambda - 1 \;=\; \frac{\Delta L_1 - \Delta E}{L_1 - E} \;=\; \frac{\omega_1 \Delta L_1 - q\Delta C_1}{\omega_1 L_1 - qC_1} \;=\; \frac{\rho \Delta C_1 + \Delta C_2}{\rho C_1 + C_2} . \; \square$$

Relying on the first-order condition $\omega_2 H' L_2 = \pi$ allows one to express the ability rent Y as a strictly proportional function of E:

$$Y(\omega_2, \pi, L_2) \;=\; \omega_2 H(E)L_2 - \pi E \;=\; H\frac{\pi}{H'} - \pi E \;=\; (\frac{1}{\eta} - 1)\pi E . \tag{12}$$

As a result,

$$\frac{Y}{Y + \pi E} = 1 - \eta \ . \tag{13}$$

By relying on the definition of the expenditure function and by invoking Hotelling's lemma one obtains

$$q[\rho C_{1x} + C_{2x}] = \rho \omega_1 L_{1x} + \omega_2 H L_{2x} \quad \text{for} \quad x = q, \omega_1, \rho, \varphi . \tag{14}$$

The relationship (14) extends to the Δ-notation:

$$q[\rho \Delta C_1 + \Delta C_2] = \rho \omega_1 \Delta L_1 + \omega_2 H \Delta L_2 \ . \tag{15}$$

Proposition 2 can now be proved by relying on (15), Proposition 1, (3) and (13):

$$\frac{\Delta L_2}{L_2} \underset{(15)}{=} \frac{1}{\omega_2 H L_2} [q \rho \Delta C_1 + q \Delta C_2 - \rho \omega_1 \Delta L_1]$$

$$\underset{\text{Prop. 1}}{=} \frac{1}{\omega_2 H L_2} [q \rho C_1 + q C_2 - \rho \omega_1 L_1] \frac{\Delta E}{E}$$

$$\underset{(3)}{=} \frac{1}{\omega_2 H L_2} [\omega_2 H L_2 - \pi E] \frac{\Delta E}{E}$$

$$= \frac{Y}{Y + \pi E} \frac{\Delta E}{E} \underset{(13)}{=} (1 - \eta) \frac{\Delta E}{E} . \ \square$$

The *proof of Proposition* 3 relies on isoelasticity of the earnings function, $H(E) = hE^{\eta}$, $\eta < 1$. Together with (12) isoelasticity implies

$$\varepsilon_{Y/x} = \varepsilon_{E/x} + \varepsilon_{\pi/x} \quad \text{for} \quad x = q, \rho, \omega_1, \varphi . \tag{16}$$

where $\varepsilon_{X/x} = \frac{x}{X} \frac{\partial X}{\partial x}$ denotes the elasticity of X with respect to x.

Remark: Assuming $H(E) = hE^{\eta}$, $\eta < 1$, one obtains

$$\frac{\Delta L_2}{L_2} = \frac{Y}{Y + \pi E} \frac{\Delta E}{E} + \frac{\pi - p}{\pi} .$$

Proof: By the definition of the Δ-operator in (7) one obtains

$$[Y + \pi E]\frac{\Delta L_2}{L_2} = \omega_2 H \Delta L_2$$

$$= (q-1)\omega_2 H L_{2q} + \frac{1}{\rho}[\frac{q-1}{q}(\rho - r)\omega_1 + r(\omega_1 - w_1)]\omega_2 H L_{21}$$

$$+ \frac{\rho - r}{q}\omega_2 H L_{2\rho} + \frac{1}{\rho}[\frac{q-1}{q}(\rho - r)\varphi + r(\varphi - f)]\omega_2 H L_{2\varphi}$$

$$= (q-1)Y_q + \frac{1}{\rho}[\frac{q-1}{q}(\rho - r)\omega_1 + r(\omega_1 - w_1)][Y_1 + \rho E]$$

$$+ \frac{\rho - r}{q}[Y_\rho + (\omega_1 + \varphi)E] + \frac{1}{\rho}[\frac{q-1}{q}(\rho - r)\varphi + r(\varphi - f)][Y_\varphi + \rho E]$$

$$\underset{(16)}{=} (q-1)\frac{Y}{E}E_q + \frac{1}{\rho}[\frac{q-1}{q}(\rho - r)\omega_1 + r(\omega_1 - w_1)]\frac{Y}{E}E_1$$

$$+ \frac{\rho - r}{q}[\frac{Y}{E}E_\rho + \frac{Y}{\rho} + (\omega_1 + \varphi)E]$$

$$+ \frac{1}{\rho}[\frac{q-1}{q}(\rho - r)\varphi + r(\varphi - f)]\frac{Y}{E}E_\varphi$$

$$+ [\frac{q-1}{q}(\rho - r)(\omega_1 + \varphi) + r(\omega_1 + \varphi - w_1 - f)][\frac{Y}{\pi} + E]$$

$$= \frac{Y}{E}\Delta E + \frac{\pi - p}{\pi}[Y + \pi E]. \ \square$$

Proposition 3 is proved when comparing the Remark with Proposition 2 and when making use of Proposition 1 and (13). $\square$

References

Arrow, K.J., 1962, The economic implications of learning by doing, Review of Economic Studies 29, 155−173.

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.

Bovenberg, A.L. and B. Jacobs, 2001, Redistribution and education subsidies are Siamese twins, CEPR Disc. Paper No. 3099.

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

Carneiro, P. and J.J. Heckman, 2003, Human capital policy, in: Inequality in America, J.J. Heckman and A.B. Krueger, eds., MIT Press, 77−239.

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.

Heckman, J.J., 1976, A life-cycle model of earnings, learning, and consumption, Journal of Political Economy 84, S11-S44.

Jacobs, B. and A.L. Bovenberg, 2005, Human capital and optimal positive taxation of capital income, Tinbergen Institute, Disc. Paper TI 2005-035/3.

Jacobs, B. and A.L. Bovenberg, 2007, Generalized results on optimal taxation and education policy and conditions for weak efficiency in human capital production", mimeo: University of Amsterdam/Tilburg

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.

Nielsen, S.B. and P.B. Sörensen, 1997, On the optimality of the Nordic system of dual income taxation, Journal of Public Economics 63, 311–329.

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.

Ritter, F.E., and Schooler, L.J., 2001, The learning curve, in: International Encyclopedia of the Social and Behavioral Sciences, W. Kintch, N. Smelser, and P. Baltes, eds., Oxford, 8602−8605.

Trostel, P.A., 1993, The effect of taxation on human capital, Journal of Political Economy 101, 327–350.

Trostel, P.A., 2002, Should education be publicly provided? Bulletin of Economic Research 54, 373−391.