RUHR
ECONOMIC PAPERS

Wolfram F. Richter

Taxing Education in Ramsey's Tradition

technische universität dortmund

#140

Imprint

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Prof. Dr. Thomas K. Bauer
RUB, Department of Economics, Empirical Economics
Phone: +49 (0) 234/3 22 83 41, e-mail: thomas.bauer@rub.de

Prof. Dr. Wolfgang Leininger
Technische Universität Dortmund, Department of Economic and Social Sciences
Economics – Microeconomics
Phone: +49 (0) 231/7 55-3297, email: W.Leininger@wiso.uni-dortmund.de

Prof. Dr. Volker Clausen
University of Duisburg-Essen, Department of Economics
International Economics
Phone: +49 (0) 201/1 83-3655, e-mail: vclausen@vwl.uni-due.de

Prof. Dr. Christoph M. Schmidt
RWI, Phone: +49 (0) 201/81 49-227, e-mail: christoph.schmidt@rwi-essen.de

Editorial Office

Joachim Schmidt
RWI, Phone: +49 (0) 201/81 49-292, e-mail: joachim.schmidt@rwi-essen.de

Ruhr Economic Papers #140

Responsible Editor: Wolfgang Leininger

ISSN 1864-4872 (online) – ISBN 978-3-86788-158-6

Ruhr Economic Papers #140

Wolfram F. Richter

Taxing Education in Ramsey's Tradition

technische universität dortmund

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http//dnb.ddb.de* abrufbar.

ISSN 1864-4872 (online)
ISBN 978-3-86788-158-6

Wolfram F. Richter[1]

Taxing Education in Ramsey's Tradition

Abstract

Assuming a two-period model with endogenous choices of labour, education, and saving, it is shown to be second-best efficient to deviate from Ramsey's Rule and to distort qualified labour less than nonqualified labour. Furthermore, if the earnings function displays constant elasticity, the choice of education should not be distorted. With the necessary qualifications the results extend to the case when taxpayers are heterogeneous and when the planner trades off efficiency against equity.

JEL Classification: H21, I28, J24

Keywords: Endogenous choice of education, labour, and saving; second-best efficient taxation; linear instruments; finite periods; Ramsey's Rule; Power Law of Learning

September 2009

1 *TU Dortmund University, CESifo Munich, and IZA Bonn.* – *Forthcoming in:* Journal of Public Economics. *Valuable comments and suggestions made by Christoph Braun, Thiess Büttner, Bas Jacobs, Josef F. Krems, Etienne Lehmann, Dirk Schindler, Kerstin Schneider, Klaus Wiener, and two anonymous referees are gratefully acknowledged. The comments referred in part to earlier versions of this paper.* *– All correspondence to Wolfram F. Richter, TU Dortmund, Department of Economics, 44221 Dortmund, Germany, email: Wolfram.Richter@tu-dortmund.de.*

1. Introduction

Human-capital accumulation is expected to be the driving engine of economic growth and development in the 21st century. The setting of correct incentives for education must therefore rank high on the political agenda. Unfortunately, the economic understanding of optimal education policy is still rather limited. A major reason is that education is a highly complex process, which is affected both by taxation and by potential market failures. This paper focuses on taxation and on the effects it has on the most basic trade-offs in education. As a consequence, the model studied is simple. Taxpayers have to make a static decision on education, saving, and qualified and nonqualified labour. Such a selection of endogenous choices can be justified as follows. Education raises the productivity of labour, which makes it necessary to differentiate between qualified and nonqualified labour. Education takes time and hence causes a cost in the forgone income earned by nonqualified labour. One would however not talk of education if forgone earnings were the sole cost of education. The use of the term *education* indicates that there are educators instructing the learners and these educators must be paid. This suggests differentiating between the opportunity cost of learning and the monetary cost of education.[1] Finally, education has features of an investment activity. Its costs are only justified if the return can keep abreast of alternative investments. Hence saving must be modelled along with education.

The model fulfilling such requirements is a straightforward extension of the standard two-period life-cycle model, and the analysis of optimal taxation follows Ramsey's tradition. The paper starts by focussing on a representative taxpayer in Sections 2 and 3. Extensions to heterogeneous taxpayers are derived in Section 4. As a first major result it is shown to be second-best efficient to deviate from Ramsey's Rule and to distort qualified labour less than nonqualified labour (Proposition 1). No similar result is known from the Mirrlees approach to the optimal taxation of education. The result holds for arbitrary utility and learning functions. The efficient reduction of nonqualified labour equals that of education and consumption in relative terms. With the General Theory of Second Best (Lipsey and Lancaster, 1956/57) in mind, one might think it efficient to spread tax distortions uniformly across all feasible margins. There are, however, particular scenarios where such an inference is unwarranted. As others have shown before, there are well-selected utility functions for which it is second best not to distort saving and it is equally second best not to distort education if the learning function is isoelastic. For the sake of brevity the latter is called the Education Efficiency Proposition. First versions have been proved by Bovenberg and Jacobs (2005), and Propositions 2 and 3 are variations designed to clarify the assumptions needed to prove the Proposition. It is shown that the assumptions made by Jacobs and Bovenberg (2008) in their latest version for heterogeneous taxpayers can be relaxed in the Ramsey framework (Proposition 3). The most critical assumption needed to prove the Proposition is that the elasticity of learning must be constant across individuals and varying choices of education. This assumption will be defended by referring to the cognitive psychology literature, which provides impressive empirical evidence in favour of such constancy if only the learning program is kept fixed. This phenomenon is known as the Power Law of Learning. The suggested policy implication is to ensure undistorted educational choices within particular learning programs (*intensive margin*). Whether and when it is optimal to distort the choice between competing learning programs (*extensive margin*) is a question leading beyond the scope of the present study.

As mentioned, the paper focuses only on the effects taxation has on the basic trade-offs in education. Such an objective necessitates ignoring various extensions and complications, which have been the subject of scrutiny in the literature. Thus credibility problems of government policy will be ruled out. The possible time inconsistency of education policy is studied by Boadway, Marceau, and Marchand (1996) and Andersson and Konrad (2003). The return to education will be considered to be certain. Uncertainty is addressed by da Costa and Maestri (2007) and Anderberg (2008). Informational asymmetry and availability of nonlinear tax instruments will be ruled out. The so-called Mirrlees approach to optimal taxation is followed by

[1] The importance of such a differentiation has been stressed before by Trostel (1993, 1996). Nielsen and Sörensen fail to make it, and this strongly biases their results. See Section 3 below. Differentiation is however not crucial in the Mirrlees approach to the optimal taxation of education. See Jacobs and Bovenberg (2008).

Bovenberg and Jacobs (2005), Wigger (2004), and Jacobs and Bovenberg (2008). Finally, in contrast with Trostel (1993, 1996), Jones, Manuelli, and Rossi (1997), Atkeson, Chari, and Kehoe (1999), and Richter and Braun (2009), this paper analyses taxation in a purely static framework.

The paper is structured as follows. Section 2 sets up the model of a representative taxpayer. Section 3 analyses the structure of second-best efficient policy for the representative taxpayer. Section 4 provides extensions to heterogeneous taxpayers. Section 5 discusses connections to the literature. 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 a time $L_1 - E$ is spent in the market, while a 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 *learning function* $H(E)$ displays positive but diminishing returns, $H'>0>H''$. It is suggestive to interpret L_2 and HL_2 as *qualified labour* and *effective qualified labour*, respectively. Likewise, we will use the terms *nonqualified labour* and *nonqualified non-leisure* for the quantities $L_1 - E$ and L_1, respectively. Education has an opportunity cost in forgone earnings captured by $\omega_1 E$. This *cost of learning* adds to the (*monetary*) *cost of education*, for which college fees may stand. For the sake of simplicity the monetary cost is likewise modelled as a linear function of the amount of education, φE. The share of first-period income that is spent neither on education nor on consumption is saved:

$$S = \omega_1 (L_1 - E) - \varphi E - C_1 = \omega_1 L_1 - (\omega_1 + \varphi) E - C_1 . \qquad (1)$$

By way of normalization, the price of consumption is set equal to one. The gross rate of return to saving is denoted by ρ. Second-period consumption is constrained by income earned:

$$C_2 = \rho S + \omega_2 H(E) L_2 \ . \tag{2}$$

All prices are after taxes and subsidies, and the question is which combination of taxes and subsidies is constrained 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

$$\rho C_1 + C_2 = \rho \omega_1 L_1 + \omega_2 H(E) L_2 - \pi E \ , \tag{3}$$

stated in second-period units. Interpret $\pi \equiv \rho(\omega_1 + \varphi)$ as the *effective* (*unit*) *cost of education*.

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 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 rent income. However, the focus is here on an endogenous choice of L_2. Hence *Y* has to be interpreted as quasi-rent income, the source of which is learning and its diminishing return. Note that *Y* is a monotone increasing, convex function of L_2:

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

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 $p = r(w_1 + f)$. Here *r* is the gross rate of return to saving before taxes and subsidies, and *f* is the (unit) cost of education before taxes and subsidies. The choice of education is *efficient* or *not distorted* if the tax wedge δ between the marginal social return and the effective social cost,

$$\delta \;\equiv\; \frac{w_2 H'(E) L_2}{r} - (w_1 + f) \;=\; \frac{\pi}{r}\Big[\frac{w_2 H' L_2}{\pi} - \frac{p}{\pi}\Big] \;=\; \frac{\pi}{r}\Big[\frac{w_2}{\omega_2} - \frac{p}{\pi}\Big] \,,$$

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

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

The taxpayer's expenditure function is defined as

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

in C_1, C_2, L_1, L_2 such that $U(C_1, C_2, L_1, L_2) \geq u$. Assume that the expenditure function is twice differentiable. This requires the minimization to be well behaved, which is only guaranteed if the concavity of U as a function of L_2 is so strong that it compensates for the convexity of Y in L_2. (See the example discussed at the end of Section 4.) Hence, the assumption of twice differentiability is less innocuous than it would be if education were exogenous.

By relying on a straightforward generalization of the textbook version of *Hotelling's lemma* one derives the identities $e_1 \equiv \frac{\partial e}{\partial \omega_1} = -\rho(L_1 - E)$, $e_2 \equiv \frac{\partial e}{\partial \omega_2} = -HL_2$, $e_\varphi = \rho E$, and $e_\rho = C_1 - \omega_1 L_1 + (\omega_1 + \varphi)E = -S$, where subscripts of e indicate partial derivatives. The capital letters L_i, S, and C_1 have to be interpreted as Hicksian supply and demand functions. This means that they have to be evaluated at $\omega_1, \omega_2, \rho, \varphi$, and u. As a result, the choice of education reads $E = E(\omega_2, \rho(\omega_1 + \varphi), L_2(\omega_1, \omega_2, \rho, \varphi; u))$ when the functional relationships are fully spelled out.

The government faces the need to raise revenue. Four linear tax instruments are available, each of which is distorting. The taxes are levied on period *i*'s labour income, on capital income, and on the monetary cost of education. For the most part of the analysis we choose to model the tax instruments implicitly as the difference between prices before and after tax. This means that the tax on period i's labour income is modelled by $w_i - \omega_i$, the tax on capital income by $r - \rho$, and the tax on the cost of education by $\varphi - f$. It goes without saying that each tax can well take on a negative value so that it is effectively a subsidy. To find out which combination of taxes and subsidies is constrained efficient is the purpose of the analysis. Government's net revenue amounts to

$$T \equiv (w_1 - \omega_1)(L_1 - E) + (\varphi - f)E + [(w_2 - \omega_2)H(E)L_2 + (r - \rho)S]/r.$$

By invoking Hotelling's lemma it can be written as

$$T = \frac{1}{\rho}(\omega_1 - w_1)e_1 + \frac{1}{\rho}(\varphi - f)e_\varphi + [(\omega_2 - w_2)e_2 + (\rho - r)e_\rho]/r. \qquad (5)$$

3. Second-best efficient policy

The planner's problem is to maximize net revenue (5) in $x = \varphi, \omega_1, \omega_2, \rho$ subject to the individual budget constraint e=0. In the Appendix it is shown that taking partial derivatives with respect to $x = \varphi, \omega_1, \omega_2, \rho$, invoking Hotelling's lemma, and eliminating the Lagrange multiplier yields the following system of three first-order conditions:

$$\frac{\Delta E}{E} = \frac{\Delta L_1 - \Delta E}{L_1 - E} = \frac{\Delta(HL_2)}{HL_2} = \frac{\Delta C_1 - \omega_1 \Delta(L_1 - E) + \varphi \Delta E}{C_1 - \omega_1(L_1 - E) + \varphi E}, \qquad (6)$$

where the total differentiation operator Δ is defined on arbitrary functions $X = X(\omega_1, \omega_2, \rho, \varphi; u)$ by

$$\Delta X \equiv \frac{1}{\rho}(\omega_1 - w_1)X_1 + \frac{1}{\rho}(\varphi - f)X_\varphi + \frac{1}{r}(\omega_2 - w_2)X_2 + \frac{\rho - r}{r}X_\rho. \quad (7)$$

According to (7), ΔX equals the weighted sum of the partial derivatives of X with the weights given by the tax wedges. It is an approximation of the total change in X when taxes are chosen efficiently. If it were efficient not to tax the opportunity cost and the monetary cost of education ($\omega_1 = w_1, \varphi = f$), then the last term in (6) could be interpreted as a relative change in saving. This is because

$$S = -[\,C_1 - \omega_1(L_1 - E) + \varphi E\,] = -\,e_\rho \text{ and}$$

$$\Delta S = -\,\Delta e_\rho = -[\,\Delta C_1 - \omega_1 \Delta(L_1 - E) + \varphi \Delta E\,] + \frac{1}{\rho}[(\omega_1 - w_1)(L_1 - E) - (\varphi - f)E].$$

By relying on some simple algebraic manipulations and by making use of

$$\frac{\Delta(HL_2)}{HL_2} = \frac{\Delta L_2}{L_2} + \frac{\Delta H}{H} = \frac{\Delta L_2}{L_2} + \eta \frac{\Delta E}{E}, \tag{8}$$

where the elasticity $\eta \equiv EH'/H$ may well be non-constant in *E*, (6) can be restated as

$$\frac{\Delta E}{E} = \frac{\Delta L_1}{L_1} = \frac{\Delta C_1}{C_1} \quad \text{and} \quad \frac{\Delta L_2}{L_2} = (1-\eta)\frac{\Delta E}{E}. \tag{6′}$$

As differentiation is additive, (6′) could equivalently be written in the form where the ratio $\Delta L_1 / L_1$ is replaced with the ratio $\Delta(L_1 - E)/(L_1 - E)$. In the Appendix (6′) is shown to imply

Remark 1: $\dfrac{\Delta C_2}{C_2} = \dfrac{\Delta E}{E}$.

Hence the quantities C_1, C_2, L_1, *E*, $L_1 - E$, and HL_2 should be reduced in the same proportion from their pre-tax values, whereas L_2 should be reduced to a lesser degree, when all these demand and supply functions are interpreted in the Hicksian sense. The equiproportionate reduction is something one would clearly expect in view of Ramsey's (1927) characterization of efficient taxation. The striking result concerns L_2. Obviously, efficiency requires reducing qualified labour relatively less than non-qualified labour. The ratio equals $1-\eta$, and it decreases in η. In other words, the more elastic the individual learning function is, the less should qualified labour be reduced in relative terms. Although this makes good sense, one must note that it fails to agree with Ramsey's Rule of reducing *all* household choices equiproportionately. Only the effective labour HL_2 is reduced equiproportionately. As *H*=*H*(*E*) reacts elastically, L_2 is reduced less.

Proposition 1: Second-best efficient policy requires reducing

(i) education, consumption, nonqualified non-leisure and labour, and effective qualified labour equiproportionately while reducing

(ii) qualified labour to a lesser degree in accordance with

$$\frac{\Delta L_2}{L_2} = (1-\eta)\,\frac{\Delta L_1}{L_1} \ . \qquad (9)$$

Proposition 1 raises the question as to which choices of $\varphi, \omega_1, \omega_2$, and ρ (and the associated tax rates) are second best. Clearly, one cannot expect any interesting relationship to hold in full generality. Instead, one has to assume special functions *U* or *H*. By varying these assumptions different characterizations of second-best tax rates are obtained. Some of these are more novel than others as compared with standard Ramsey tax results. An example of little novelty concerns saving. As shown by Atkinson and Stiglitz (1972) and Sandmo (1974), saving should not be taxed if utility is weakly separable in consumption and labour and homothetic in consumption. This result extends to the present framework if *U* is specified as $U(G(C_1, C_2), L_1, L_2)$ with some homogeneous function *G*. See the earlier draft of this paper (Richter, 2008). In what follows the focus is on characterizations of second-best tax rates that contrast with standard Ramsey results. Two such characterizations are derived. The first one is on education, and the second one is on labour.

On applying the differentiation operator Δ to the first-order condition associated with education, $H' L_2 = \pi / \omega_2$, one obtains

$$\frac{EH''}{H'}\frac{\Delta E}{E} + \frac{\Delta L_2}{L_2} = \frac{\Delta(\pi/\omega_2)}{\pi/\omega_2}$$

$$= \frac{1}{\pi}[(\omega_1 - w_1) + (\varphi - f) - \frac{1}{r}(\omega_2 - w_2)\frac{\pi}{\omega_2} + \frac{\rho - r}{r}(\omega_1 + \varphi)]$$

$$= \frac{1}{r}[\frac{w_2}{\omega_2} - \frac{p}{\pi}]. \qquad (10)$$

By relying on (6′) and substituting for $\Delta L_2 / L_2$ (10) implies

$$\eta' \frac{H}{H'} \frac{\Delta E}{E} = [1 + \frac{EH''}{H'} - \frac{EH'}{H}] \frac{\Delta E}{E} = \frac{1}{r}[\frac{w_2}{\omega_2} - \frac{p}{\pi}] \ . \qquad (11)$$

Obviously, the right-hand side must vanish if η'=0.

Proposition 2*:* If, and only if the individual learning function is isoelastic, it is efficient not to distort the choice of education.

Proposition 2 is only the first version of the *Education Efficiency Proposition* derived in this paper. An intuitive explanation is the following. Second-best efficient policy needs to be sustained by private maximization behaviour. This is however ensured only if condition (10) holds. If the right-hand side of (10) does not vanish, a discrepancy results between maximizing private and social ability rents, $\omega_2 HL_2 - \pi E$ and $w_2 HL_2 - pE$. This is so because the last term on the right-hand side of the following identity does not vanish:

$$w_2 HL_2 - pE = \frac{w_2}{\omega_2}[\omega_2 HL_2 - \pi E] + \pi[\frac{w_2}{\omega_2} - \frac{p}{\pi}]E \ .$$

Hence maximizing efficiency does not require maximizing the social ability rent. In this case, the planner trades off the objective of maximizing the social ability rent against the objective of minimizing the efficiency loss resulting from distorted choices of the utility-generating quantities C_1, C_2, L_1, and L_2. These two objectives are only separable if (4) holds, i.e. if the right-hand side of (10) vanishes. Compatibility of (10) with (6′) however requires that (4) holds if, only if the elasticity of the individual learning function is constant.

If some positive tax revenue is to be generated, second-best efficient policy calls for reducing education, ΔE<0. Condition (11) then implies η'>0 if, and only if $w_2/\omega_2 < p/\pi \Leftrightarrow w_2/p < \omega_2/\pi$. Hence, education should be effectively subsidized if, and only if η is strictly increasing at the second-best efficient level of *E*.

Remark 2: If, and only if the elasticity of the individual learning function is strictly increasing at the second-best level of education, should education be effectively subsidized.

Combining Propositions 1 and 2 implies that efficient policy well tolerates a reduction in education. This reduction cannot be interpreted, however, as a (conditional) distortion of education. This observation allows one to qualify Trostel (1993), who stresses the negative effect of proportional income and consumption taxation on education. To make the point clear, consider some proportional tax on labour income and allow monetary costs of education to be tax-deductible. In this case w_2 is reduced in the same proportion as *p*. As a result, all individual choices of C_1, C_2, L_1, and L_2 will be distorted. Still, the partial efficiency condition (4) holds by construction.

To illustrate the effect of endogenous education on efficient labour taxation, consider the scenario given by

$$H = E^{\eta}, \quad U = G(C_1, C_2) - V_1(L_1) - V_2(L_2), \quad \nu_i \equiv L_i V_i'' / V_i' \ (i=1,2), \quad (12)$$

and linear homogeneous *G*. As mentioned before, the taxpayer's optimization is only well behaved if the concavity of *U* as a function of L_2 is strong enough to compensate for the convexity of the ability rent *Y* in L_2. In terms of (12) this means that $\nu_2 > \eta/(1-\eta)$ has to hold by assumption. Define taxes τ_i in exclusive form by setting $w_2 \equiv (1+\tau_2)\omega_2$ and $w_1 \equiv (1+\tau_1)\omega_1$. In the Appendix it is shown that wage taxes are second best if they satisfy the condition

$$\frac{\tau_2}{\tau_1} = \frac{(1-\eta)\nu_2 - \eta}{\nu_1}. \quad (13)$$

As $\nu_2 > \eta/(1-\eta)$ is to hold by assumption, the numerator on the right-hand side of (13) is positive. For $\eta = 0$, (13) is the familiar *Inverse Elasticity Rule*. According to this rule, wage taxes τ_i should be set inversely proportional to the wage elasticities of labour supplies, $1/\nu_i$. This rule is extended by (13) to cope with endogenous education. The effect of education is to reduce τ_2 relative to τ_1. Just note that $(1-\eta)v_2 - \eta < v_2$.

4. Extensions

Propositions 1 and 2 are clear-cut results, and it is natural to ask whether they continue to hold in more general settings. Two extensions are of particular interest. One allows for endogeneity of factor prices, and the other allows for heterogeneity of taxpayers. As the former extension implies no surprising result, it will only be sketched.

The claim is that Propositions 1 and 2 go through when prices are endogenous. In fact, it has been shown in the earlier draft of this paper (Richter, 2008) that Propositions 1 and 2 continue to hold if the second-period factor prices r and w_2 are endogenous. Two assumptions must however hold. The first one states that the government must be able to issue debt B, so that only the excess of savings over public debt, $K=S-B$, is a factor of production. The second assumption requires constant returns to scale of production in effective qualified labour HL_2 and capital K. Such constant returns make the Diamond–Mirrlees (1971) Production Efficiency Theorem applicable, and it is second best not to distort production but only consumption.

A more interesting extension concerns heterogeneity of taxpayers. As will be shown next, the Education Efficiency Proposition holds even if the planner trades off efficiency against equity and even if the set of policy instruments is incomplete. However, it must be assumed that the right policy instruments are available and that the elasticity of learning is constant, both in the level of education and across individuals. To be more precise, let $n = 1,..,N$ be the parameter identifying a particular taxpayer. Taxpayers are assumed to differ in preferences and the productivity of learning, but not in the elasticity of learning. Hence $u_n = U^n(C_1, C_2, L_1, L_2)$ and $H^n(E) = h_n E^\eta$. Let E_n, L_{2n}, etc. be the choices made by n, and let T^n denote the taxes paid by n on labour income, savings, and the cost of education as specified by (5). In order to model redistribution, assume that n receives some exogenous income g_n financed on the margin by taxing qualified labour income and the cost of education.

Hence φ and ω_2 belong to the set of available policy instruments. The planner then maximizes net aggregate tax revenue subject to the constraints that individual budgets are balanced and that welfare W remains constant:

$$\sum_n [T^n - g_n / r] \text{ in } \varphi, \omega_2, u_n \text{ subject to} \tag{14}$$

$$g_n = e(\omega_1, \omega_2, \rho, \varphi; u_n) \quad , \quad (\lambda_n) \tag{15}$$

$$\text{and } W(u_1,..,u_N) = \text{constant.} \tag{16}$$

λ_n denotes a Lagrange variable. In the Appendix the following result is derived.

Proposition 3: Assuming heterogeneous taxpayers but constancy of the learning elasticity η in n, E, and assuming availability of φ, ω_2, it is optimal not to distort education.

Proposition 3 generalizes a version of the Education Efficiency Proposition for heterogeneous taxpayers that has been derived before by Jacobs and Bovenberg (2008). The generalization lies in showing that two assumptions on which the analysis of Jacobs and Bovenberg is based can be dropped. One is the availability of a poll tax, and the other is the assumption of identical utility functions.

The assumption that both φ and ω_2 are available policy instruments is noteworthy in that it implies that φ cannot simply be substituted by ω_1 without affecting efficiency. This means that it is not irrelevant in the Ramsey framework which kind of educational cost is taxed or subsidized. Subsidizing the monetary cost of education is not equivalent to subsidizing the opportunity cost of forgone earnings, as shown in more detail in the earlier draft of the present paper (Richter, 2008). A subsidy to the monetary cost of education targets the choice of education without directly affecting utility. This is different with ω_1. A subsidy to the cost of forgone earnings affects both the choice of education and the choice of non-qualified leisure. The lack of instrumental equivalence of φ and ω_1 has to be borne in mind when interpreting results that have been derived in models not allowing one to differentiate between the two costs of education. A case in point is Nielsen and Sörensen (1997), who analyse the merits of dual income taxation. Their main and much-cited result states that labour income should optimally be taxed progressively ($\omega_2 / w_2 < \omega_1 / w_1$) if the qualified labour supply is not too elastic and cross substitution of complementarity effects is not too strong. The present analysis suggests that this result should be interpreted with great caution because it relies on an analysis not modelling the monetary cost of education. It is as if φ would not be an available instrument, and the proven efficiency of progressive labour taxation can be attributed at most third-best status.

The constancy of the learning elasticity η in n, E is critical for Proposition 3 to hold, and it needs to be commented on. Jacobs and Bovenberg (2008) find such constancy restrictive and presumably too restrictive to serve as the basis of policy recommendations. However, constancy can well be defended by referring to the *Power Law of Learning* known from cognitive psychology. The content of this 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. The elasticity of completion is not only a constant function of practice; it also seems to be fairly constant across individuals. In any case, individual learning functions seem to differ less in their elasticity than in their level (Anderson, 2005 (1980), Chap. 6; Crossman, 1959). If practice is denoted by E and the inverse completion time of subject n by H^n, this evidence suggests specifying $H^n(E)$ as $h_n E^\eta$. The only drawback is that elasticities can differ strongly between different learning programs. This suggests relying on the Power Law of Learning if the focus is on a particular learning program and rejecting it else. The policy implication would be to ensure that educational choices are at least not distorted within particular learning programs (intensive margin). Whether and when it is optimal to distort the choice between competing learning programs (extensive margin) cannot be answered by the present study.

According to Ritter et al. (2001) “the power law of practice is ubiquitous”. Still, little reference to it can be found in the economics literature. A well-known exception is Arrow (1962). However, in Arrow’s model the learning function 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.

If many taxpayers are involved, Ramsey’s result of equiproportionate reductions in demands and supplies is known to only hold with some qualification. The reductions have to be differentiated in accordance with equity concerns. This has been shown by Diamond and Mirrlees, and a textbook version can be found in Myles (1995). Second-best efficient reductions in aggregate compensated demands are the smaller the more the goods are consumed by taxpayers whose marginal utility of income is highly valued in social terms. This result carries over to the present context but only for first-period consumption, nonqualified non-leisure and labour, education, and effective qualified labour. For a proof see the Appendix.

Proposition 4: Maximizing (14) in $\varphi, \omega_1, \omega_2, \rho, u_n$ subject to (15) and (16) implies

$$\frac{\sum_n \Delta X_n}{\sum_n X_n} = -[\frac{1}{r} - \sum_n \frac{\lambda_n X_n}{\sum_m X_m}] \tag{17}$$

for $X = C_1, L_1, L_1 - E, E, HL_2$.

The left-hand side of (17) has to be interpreted as a proportional reduction in aggregate compensated demand or supply of *X*. The right-hand side determines the size of the reduction. The reduction would be the same for all *X* if the social marginal utility of income λ_n were constant across taxpayers. If, however, such constancy fails to hold for reasons of equity, the demands and supplies of taxpayers with high values of λ_n are more highly weighted than the demands of other taxpayers, and the reduction of X is adapted accordingly.

Statement (ii) of Proposition 1 extends to heterogeneous taxpayers, with two qualifications. The first one is that the change in aggregate qualified labour is compared with the change in aggregate education. Assuming the latter change to be negative ($\sum E_n$ <0), (18) holds as stated with ">". The interpretation is that aggregate qualified labour is reduced less than the extended Ramsey rule (17) suggests. The second qualification is that productivity H^n enters as a weighting scheme. More precisely, assuming $\sum E_n$ <0, it is shown in the Appendix that the weighted proportional reduction of aggregate qualified labour is less than (17) suggests:

$$\frac{\sum_n H^n \Delta L_{2n}}{\sum_n H^n L_{2n}} > -[\frac{1}{r} - \sum_n \frac{\lambda_n H^n L_{2n}}{\sum_m H^m L_{2m}}]. \tag{18}$$

5. The Mirrlees approach to optimal taxation of education

There have been attempts by Bovenberg and Jacobs (2005) and by Wigger (2004) to characterize optimal incentives for education when adopting the modelling tradition of Mirrlees (1971) and Atkinson and Stiglitz (1976). The specific feature of this approach is the assumption of asymmetric information. In terms of the present notation it is as if the qualified wage rate $w_2(n)$ and the qualified labour supply L_{2n} are private and not public information. The planner can only verify the product of the two. In a model with education the question arises whether and to what extent the amount of education should be verifiable. Bovenberg and Jacobs study the scenario when E_n is verifiable as well as the scenario when E_n is imperfectly verifiable. The following discussion assumes that all individual choices $C_{1n}, C_{2n}, L_{1n}, E_n$, except that of the qualified labour supply, are verifiable. Full verifiability of E_n convinces to the extent that education can be measured by the years spent in institutions of education. Jacobs and Bovenberg (2008) demonstrate that it is optimal not to distort education if three assumptions hold:

(i) The planner must be able to levy a nonlinear tax $\mathcal{T}$ on qualified labour income and also to subsidize costs of education by some nonlinear scheme $\mathcal{S}$. (ii) Utility functions must be weakly separable in qualified labour and all other individual choice variables. In the present notation this means $U = U(V(C_1, C_2, L_1), L_2)$. (iii) Qualified labour income before tax, Z_2, must be weakly separable in n and L_2, on the one hand, and in education E, on the other hand, so that Z_2 can be written as $Z_2(w_2(n, L_2), E)$. Given this set of assumptions, it is optimal to equalize marginal rates of taxation and subsidization: $\mathcal{T}' = \mathcal{S}'$. As a result, not only education is undistorted, but also saving and nonqualified labour supply. The most direct way of implementing such an optimal tax-transfer system would be the following: (i) Only qualified labour income is taxed. (ii) Taxpayers are allowed to carry forward costs of education and learning and to deduct them against qualified labour income Z_2. Notice that not only monetary costs of education should be tax-deductible, but also costs of forgone earnings. See also Trostel (1996, 1993), who argues in favour of deductions even exceeding 100%.

da Costa and Maestri (2007) and Anderberg (2008) extend the analysis of optimal education policy by incorporating uncertainty. Anderberg sets up a model in which qualified labour income can be written in multiplicative form, $Z_2 = w_2(n, E_2)L_2$, and in which n takes the role of a productivity shock hitting the representative taxpayer. He demonstrates that education should not be distorted if the elasticity of Z_2 with respect to E is constant in n. The simplest specification ensuring such constancy is multiplicative: $Z_2 = w_2(n)H(E)L_2$.

It may be of interest to note that the Education Efficiency Proposition can be derived in both frameworks of Ramsey and Mirrlees but that the required assumptions differ strongly. In particular, constant elasticity of earnings with respect to education is only needed in the Ramsey setting. The explanation is as follows. In the Ramsey approach the planner has two independent efficiency objectives. One is the minimization of distortions in the quantities from which utility is derived. The other objective is the maximization of the quasi-rent income generated by education. These two optimizations are separable, given the linearity constraint on policy instruments, only if the earnings function is isoelastic in education. Separability implies that it is efficient not to distort education and to minimize the distortions in consumption and labour choices. In the cited papers standing in the Mirrlees tradition the planner has only one objective to pursue. This objective is to keep highly productive taxpayers from mimicking less productive taxpayers. For this purpose it suffices to tax qualified labour, as this is the only choice variable for which private information is assumed. As all other individual choices are observable and weakly separable from qualified labour, they should remain undistorted. Conditional efficiency of education can be sustained because the functional specification of subsidization is *a priori* not constrained and the marginal subsidy to education can therefore be set equal to the marginal tax on qualified labour income.

6. Summary

Economists are only beginning to understand the optimal setting of tax incentives for education. A major breakthrough is by Bovenberg and Jacobs (2005). The present paper contributes to the literature by analysing efficient taxation of education in Ramsey's tradition. It does so by relying on the standard two-period life-cycle model of a representative household with endogenous consumption, labour, and education. A first notable result (Proposition 1) states that Ramsey's Rule does apply to education, consumption, and nonqualified labour but not to qualified labour. Qualified labour supply should be reduced less than the other demands and supplies in relative terms. No particularly selected utility functions are needed to derive this result. The modelling strategy, however, seems to be critical. At least no similar result has been derived before within the Mirrlees framework of asymmetric information.

The drawback of the Ramsey approach is that efficient reductions of demands and supplies cannot be translated one to one into efficient tax rates. Statements about efficient tax rates are only possible if specific assumptions are made. A well-known example is the familiar Inverse Elasticity Rule. In Section 3 it is shown how this rule has to be adapted if applied to qualified and nonqualified labour. Another example is the result by Atkinson and Stiglitz (1972) and others that saving should not be distorted if utility is separable in consumption and labour and homothetic in consumption. This result directly extends to the present framework with endogenous education. While it depends on the utility function whether saving should be taxed or not, the efficiency of not distorting education only depends on the learning function. More precisely, it is shown to be efficient not to distort education if the learning function is isoelastic (Proposition 2). This result is called the Education Efficiency Proposition, and Section 4 offers an extension to many heterogeneous taxpayers. Proposition 3 generalizes earlier versions of the Education Efficiency Proposition derived by Bovenberg and Jacobs (2005) and Jacobs and Bovenberg (2008) by adopting the Mirrlees approach to optimal taxation. Two sets of assumptions must hold to prove the present version: (i) It must be possible to tax qualified labour and to tax/subsidize the monetary cost of education. (ii) The learning function must be isoelastic, and the elasticity must be constant across individuals. Implicit is the assumption that education is weakly separable from labour and ability in earning income.

It is natural to ask whether and to what extent the Education Efficiency Proposition can offer guidance in education policy. Jacobs and Bovenberg (2008) express scepticism. They do so by questioning the empirical relevance of weak separability of education from labour and ability in earning income, and even more the relevance of constant elasticity of education. In Section 4 a more positive view is suggested. It is argued that cognitive psychology provides impressive evidence for learning functions the elasticity of which does not vary either in the amount of learning or between individuals. Applicability of this Power Law of Learning is only limited by the observation that the elasticities can differ strongly between different learning programs. The policy conclusion would be that educational choices should at least not be distorted at the intensive margin. Things may be very different at the extensive margin, and not only for the reason that education fails to be weakly separable from labour and ability in earning income or that the planner trades off efficiency against equity. Dynamic complementarities may provide another strong reason to distort educational choices systematically. This point is elaborated by Richter and Braun (2009). By working with an overlapping-generations model with endogenous growth, they show that it may well be second-best efficient to subsidize education relative to the first-best level even though separability is ensured and even though efficiency is the planner's sole objective. The reason for systematic subsidization is a human-capital investment function displaying a strictly increasing elasticity of education. It is argued that in a dynamic framework such increasing elasticity may be a more appealing assumption than constant elasticity. In view of Remark 2 above it is not surprising to learn that subsidizing human-capital investment is then efficiency enhancing.

9. Appendix

The *proof of* (6) relies on taking partial derivatives of the Lagrange function $T-\lambda e$ with respect to φ, ρ, ω_1, and ω_2:

$$\frac{\partial}{\partial \varphi}[T-\lambda e]=0 \quad \Leftrightarrow \quad (\lambda-\frac{1}{\rho})e_{\varphi}=\Delta e_{\varphi}. \tag{19}$$

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

$$e_{\varphi}=\rho E \quad \text{and} \quad \Delta e_{\varphi}=\Delta(\rho E)=\rho\Delta E+\frac{\rho-r}{r}E. \tag{20}$$

Plugging (20) into (19) yields $\lambda-1/r=\Delta E/E$. Similarly one derives

$$\lambda-\frac{1}{r} \;=\; \frac{\Delta L_1-\Delta E}{L_1-E} \;=\; \frac{\Delta(HL_2)}{HL_2} \;=\; \frac{\Delta C_1-\omega_1\Delta(L_1-E)+\varphi\Delta E}{C_1-\omega_1(L_1-E)+\varphi E}.\square$$

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

$$\rho C_{1x}+C_{2x} \;=\; \rho\omega_1 L_{1x}+\omega_2 HL_{2x} \quad \text{for} \;\; x=\omega_1,\omega_2,\rho,\varphi. \tag{21}$$

The relationship (21) extends to the Δ-notation:

$$\rho\Delta C_1+\Delta C_2 \;=\; \rho\omega_1\Delta L_1+\omega_2 H\Delta L_2 \;. \tag{22}$$

Remark 1 is now easily proved by relying on (22), (6′), and (3):

$$\frac{\Delta C_2}{C_2} \underset{(22)}{=} \frac{\rho\omega_1\Delta L_1+\omega_2 H\Delta L_2-\rho\Delta C_1}{C_2}$$

$$= \frac{\rho\omega_1 L_1}{C_2}\frac{\Delta L_1}{L_1} + \frac{\omega_2 H L_2}{C_2}\frac{\Delta L_2}{L_2} - \frac{\rho C_1}{C_2}\frac{\Delta C_1}{C_1}$$

$$\underset{(6')}{=} \frac{1}{C_2}\frac{\Delta E}{E}[\rho\omega_1 L_1 + (1-\eta)\omega_2 H L_2 - \rho C_1]$$

$$\underset{(3)}{=} \frac{1}{C_2}\frac{\Delta E}{E}[C_2 + \pi E - \eta\omega_2 H L_2]$$

$$= \frac{1}{C_2}\frac{\Delta E}{E}[C_2 + (\pi - \omega_2 H' L_2)E] = \frac{\Delta E}{E}. \square$$

Proof of (13): As $G(C_1, C_2)$ is linear homogeneous, it is efficient to set $\rho = r$. Optimizing utility in consumption yields $G(C_1, C_2) = G(c(r)C_2, C_2) = G(c,1)C_2$ and $rC_1 + C_2 = [rc+1]C_2$. Set $G(c,1)/[rc+1] \equiv g$. Optimizing utility in L_1 yields $V_1' = rg\omega_1$. Let $L_1 = L_1(\omega_1)$ be the nonqualified labour supply. By definition of Δ,

$$\frac{\Delta L_1}{L_1} = \frac{1}{r}(\omega_1 - w_1)\frac{L_1'}{L_1} = -\frac{1}{r}\tau_1\frac{\omega_1 L_1'}{L_1} = -\frac{1}{r}\frac{\tau_1}{v_1}.$$

The determination of $\Delta L_2 / L_2$ is a bit more involved. The first-order condition of the taxpayer's optimal choice of L_2 is $V_2' = g\omega_2 H = g\omega_2 E^\eta$. Applying the Δ-operator and relying on (6′) yields

$$v_2\frac{\Delta L_2}{L_2} = \frac{\Delta V_2'}{V_2'} = \frac{\Delta\omega_2}{\omega_2} + \eta\frac{\Delta E}{E} \underset{(6')}{=} \frac{1}{r}\frac{\omega_2 - w_2}{\omega_2} + \frac{\eta}{1-\eta}\frac{\Delta L_2}{L_2}$$

$$= -\frac{1}{r}\tau_2 + \frac{\eta}{1-\eta}\frac{\Delta L_2}{L_2}.$$

After solving for $\Delta L_2 / L_2$ and equating $\Delta L_2 / L_2$ with $(1-\eta)\Delta L_1 / L_1$, one ends up with (13).□

The *Proof of Proposition* 3 generalizes that of (6). The first-order condition of the planner's maximization with respect to φ is obtained along the lines indicated by (19) and (20):

$$0 = \rho \sum_n [\Delta E_n - E_n(\lambda_n - \frac{1}{r})]. \tag{23}$$

The derivation with respect to ω_2 yields

$$\begin{aligned}
0 &= \sum_n [\Delta(H^n L_{2n}) - H^n L_{2n}(\lambda_n - \frac{1}{r})] \\
&= \sum_n [H^{n}{}' L_{2n}\Delta E_n + H^n \Delta L_{2n} - H^n L_{2n}(\lambda_n - \frac{1}{r})] \qquad (24)\\
&= \sum_n H^{n}{}' L_{2n}[\Delta E_n + \frac{1}{\eta} E_n \frac{\Delta L_{2n}}{L_{2n}} - \frac{1}{\eta} E_n(\lambda_n - \frac{1}{r})] \\
&\underset{(10)}{=} \sum_n \frac{\pi}{\omega_2}[\Delta E_n + \frac{1}{\eta} E_n \frac{1}{r}(\frac{w_2}{\omega_2} - \frac{p}{\pi}) + \frac{1-\eta}{\eta}\Delta E_n - \frac{1}{\eta} E_n(\lambda_n - \frac{1}{r})] \\
&\underset{(23)}{=} \frac{\pi}{\omega_2 \eta}\frac{1}{r}(\frac{w_2}{\omega_2} - \frac{p}{\pi})\sum_n E_n ,
\end{aligned}$$

from which $\frac{w_2}{\omega_2} - \frac{p}{\pi} = 0$ follows.□

Proof of Proposition 4: For *X*=*E* the statement (17) follows from dividing (23) through by $\sum E_n$. For $X = L_1 - E, HL_2$ one first has to take partial derivatives of the Lagrange function with respect to ω_1 and ω_2 yielding conditions displaying the same structure as (23):

$$\sum_n \Delta X_n = \sum_n (\lambda_n - \frac{1}{r}) X_n . \tag{25}$$

Dividing through by $\sum X_n$ yields (17) for $X = L_1 - E$ and HL_2. The case of $X = L_1$ follows from the linearity of (25) in *X*. Taking the partial derivative of the Lagrange function with respect to ρ yields

$$\sum_n [\Delta C_{1n} - \omega_1 \Delta(L_{1n} - E_n) + \varphi \Delta E_n] = \sum_n (\lambda_n - \frac{1}{r})[C_{1n} - \omega_1 (L_{1n} - E_n) - \varphi E_n] .$$

Making use of (25) for $X = L_1 - E$, E yields (25) for $X = C_1$ from which (17) follows. It is noted without proof that (25) and (17) hold for $X = C_2$ only if exogenous income g_n =0 for all n.

As to the inequality (18), assume $\sum_n \Delta E_n$ < 0. Then (24) implies

$$\sum_n \lambda_n H^n L_{2n} - \frac{1}{r} \sum_n H^n L_{2n} = \sum_n [H^{n\,\prime} L_{2n} \Delta E_n + H^n \Delta L_{2n}]$$

$$= \frac{\pi}{\omega_2} \sum_n \Delta E_n + \sum_n H^n \Delta L_{2n} < \sum_n H^n \Delta L_{2n} .$$

Dividing through by $\sum H^n L_{2n}$ yields (18).□

References

Anderberg, D., 2008, Optimal policy and the risk-properties of human capital reconsidered, mimeo: Royal Holloway University of London.

Anderson, J.R., 2005 (1980), Cognitive psychology and its implications, New York, 6th ed.

Andersson, F. and K.A. Konrad, 2003, Human capital investment and globalization in extortionary states, Journal of Public Economics 87, 1539–1555.

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.

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

Atkinson, A.B. and J.E. Stiglitz, 1976, The design of tax structure: Direct versus indirect taxation, Journal of Public Economics 6, 55–75.

Boadway, R., N. Marceau, and M. Marchand, 1996, Investment in education and the time inconsistency of redistributive tax policy, Economica 63, 171–189.

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

Crossman, E.R.F.W., 1959, A theory of the acquisition of speed-skill, Ergonomics 2, 153–166.

da Costa, C.E. and L.L. Maestri, 2007, The risk properties of human capital and the design of government policies, European Economic Review 51, 695–713.

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.

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, 1997, On the optimal taxation of capital income, Journal of Economic Theory 73, 93−117.

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

Mirrlees, J.A., 1971, An exploration in the theory of optimum income taxation, Review of Economic Statistics 38, 175–208.

Myles, 1995, Public Economics, Cambridge University Press.

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., 2008, Taxing education in Ramsey's tradition, CESifo WP No. 2586.

Richter, W.F. and C. Braun, 2009, Efficient subsidization of human capital accumulation with overlapping generations and endogenous growth, TU Dortmund University, mimeo.

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.

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

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

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

Wigger, B.U., 2004, Are higher learning subsidies second best? Scandinavian Journal of Economics 106, 65–82.