Taxing Education in Ramsey's Tradition

WOLFRAM F. RICHTER

CESIFO WORKING PAPER NO. 2586
CATEGORY 1: PUBLIC FINANCE
MARCH 2009

CESifo *Working Paper No. 2586*

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. The result holds for arbitrary utility and learning functions. Efficient incentives for education and saving are analysed under conditions of second and third best. It is argued that efficient tax policy should care more about incentives for education than for saving.

JEL Code: H21, I28, J24.

Keywords: endogenous choice of education, labour, and saving, second-best efficient taxation, Power of Law of Learning.

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

December, 2008

In Section 4 this paper draws on Richter (2007 and 2008). It replaces these earlier studies by presenting their results in consolidated form and by extending them in Sections 3, 5, and 6.

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 only on taxation and the effects taxation has on the very 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 Bovenberg and Jacobs (2005), Wigger (2004) and Jacobs and Bovenberg (2008). Finally, other than Trostel (1993 and 1996), Jones, Manuelli, and Rossi (1997) and Atkeson, Chari, and Kehoe (1999) this paper analyses taxation in a static framework.

The model studied in this paper is the most simple one can think of when analysing the basic trade-offs of optimal education policy in taxation. For the most part the focus is on a representative taxpayer although implications of taxpayer's heterogeneity are also discussed. The taxpayer has to decide on education, saving, qualified and nonqualified labour. The modelling strategy can be justified as follows. Education raises the productivity of labour. This makes it necessary to differentiate between qualified and nonqualified labour. Education takes time and hence causes a cost in foregone nonqualified labour income. One would however not talk of education if foregone earnings were the sole cost of education. The term of education suggests that there are educators instructing the learners and these educators must be paid. This suggests differentiating between (opportunity) costs of learning and (monetary) costs of education.[1] Finally, education has features of investment activity. The costs are only born 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 lifecycle model and the analysis of optimal taxation stands in Ramsey's tradition. 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. No similar result (Proposition 2) is known from the Mirrlees approach and it holds for arbitrary utility and learning functions. The efficient reduction of nonqualified labour equals the one of education and saving 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 (Proposition 4) and it is equally second-best not to distort education if the learning function is isoelastic (Proposition 3). 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 3, 5, and 7 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 7). 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. The 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. The evidence 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

[1] The importance of such a differentiation has been stressed before by Trostel, 1993 and 1996. Nielsen and Sörensen fail to differentiate and this strongly biases their results. See Section 5 below. Differentiation is

beyond the scope of the present study. The final point made by this paper is a political one: Tax policy should care more about efficient incentives for education than for saving. The argument relies on proving that efficiency in education is a more robust result than efficiency in saving when the restrictions on the set of available policy instruments are strengthened.

The paper is structured as follows. Section 2 sets up the model of a representative taxpayer. Section 3 demonstrates that it is without any loss of generality if the analysis of efficient taxation is carried through for exogenous factor prices. Section 4 is on second-best policy and Section 5 on third-best polices. Section 6 looks at the case of heterogeneous taxpayers. Section 7 discusses connections to the literature. Section 8 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 while 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 the endogenously determined wage rate and where 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. Equally we will refer to *nonqualified labour* and *nonqualified non-leisure* in the case of $L_1 - E$ 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 equally modelled as a linear function of the amount of education, φE. The share of first-period income that is neither spent on education nor on consumption is saved,

however not crucial in the Mirrlees approach to the optimal taxation of education. See Jacobs and Bovenberg

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

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,

$$C_2 = \rho S + \omega_2 H(E) L_2 \ . \qquad (2)$$

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

$$\rho C_1 + C_2 = \rho\omega_1 L_1 + \omega_2 H(E) L_2 - \pi E \qquad (3)$$

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 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 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:

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

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

---

(2008).

$p = r(w_1 + f)$. r is the gross rate of return to saving before taxes and subsidies. Equally f is the (unit) cost of education before taxes and subsidies. The choice of *education is efficient* if the tax wedge δ between the marginal social return and the effective social cost,

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

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

$$\frac{w_2}{p} = \frac{\omega_2}{\pi} . \qquad (4)$$

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

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 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 $\varphi - f$. Additionally, we allow for public debt, B. *A priori* each tax may well be negative and hence an effective subsidy. To find out which combination of taxes and subsidies is constrained efficient is the subject of the analysis. Government revenue is given by

$$T_1 \equiv (w_1 - \omega_1)(L_1 - E) + (\varphi - f)E + B$$

in the first period and by

$$T_2 \equiv (w_2 - \omega_2)H(E)L_2 + (r - \rho)S - rB$$

in second period.

3. The case of endogenous factor prices

A driving assumption of the analysis is that the individual return to learning is decreasing while the return to saving is constant. Such a differentiation is justifiable on the level of individual behaviour. It is clearly no valid description of aggregate behaviour. In a closed economy with fixed labour supply the returns to saving cannot be constant. As a result one might conjecture that any reason for taxing saving and education differently vanishes when factor prices are endogenous.[2] This section is to show that such a conjecture is not correct. All the results derived below continue to hold if second-period's factor prices r and w_2 are endogenous. In order to prove this claim the model is enriched by endogenous production F in period two. Production is linear homogeneous in capital, K, and effective labour, HL_2. Hence education is labour augmenting. By referring to capital market clearing, capital is substituted by $K = S - B$ in what follows.

The tax planner is assumed to maximize the first period's tax revenue T_1 in $\omega_1, \omega_2, w_2, r, \rho, B, \varphi$ subject to the constraints

$$T_2 = \text{constant}, \qquad (5)$$

$$0 = e, \qquad (6)$$

$$r = F_K(HL_2, S - B) \text{ and} \qquad (7)$$

[2] In fact, a referee argued this way.

$$w_2 = F_L(HL_2, S-B)\,. \tag{8}$$

(7) and (8) follow from profit maximization. (6) is the representative taxpayer's budget constraint. 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. See the example below in Section 4. By invoking Hotelling's lemma the planner's optimization can be written as

$$\max T_1 = [B - \frac{1}{\rho}(w_1 - \omega_1)e_1 - \frac{1}{\rho}(f - \varphi)e_\varphi]$$

in $\omega_1, \omega_2, w_2, r, \rho, B, \varphi$ subject to the constraints

$$T_2 = -[rB + (w_2 - \omega_2)e_2 + (r-\rho)e_\rho] = \text{constant} \qquad (\mu) \qquad (5')$$

$$0 = e \qquad (\lambda) \qquad (6')$$

$$r = F_K(-e_2, -e_\rho - B) \qquad (\alpha) \qquad (7')$$

$$w_2 = F_L(-e_2, -e_\rho - B) \qquad (\beta) \qquad (8')$$

Setting partial derivatives of the Lagrange function equal to zero yields:

$$\frac{\partial}{\partial r}: \quad \alpha = \mu(B + e_\rho) = -\mu K$$

$$\frac{\partial}{\partial w_2}: \quad \beta = \mu e_2 = -\mu HL_2$$

$$\frac{\partial}{\partial B}: \quad 1 - \mu r = -\alpha F_{KK} - \beta F_{LK} = \mu[KF_{KK} + HL_2 F_{LK}] = 0$$

where the last equality holds because of constant returns to scale. As a result, $\mu = 1/r$. For the sake of brevity, write $T \equiv T_1 + \mu T_2 = T_1 + T_2/r$. For $x = \varphi, \omega_1, \omega_2, \rho$ one obtains

$$\frac{\partial}{\partial x}: \quad \frac{dT}{dx} - \lambda e_x = -[\alpha F_{KL} + \beta F_{LL}]e_{2x} - [\alpha F_{KK} + \beta F_{LK}]e_{\rho x}$$

$$= \mu[KF_{KL} + HL_2F_{LL}]e_{2x} + \mu[KF_{KK} + HL_2F_{LK}]e_{\rho x} = 0 \qquad (9)$$

where the last equality once more holds because of linear homogeneity in production. As a result one may safely ignore the production side when determining the second-best efficient tax structure. It makes no real difference whether the tax structure is determined for the closed or the small open economy.

Proposition 1: If returns to scale are constant, the efficient tax structure is independent of the production side. The efficient tax structure in the closed economy is the same as in the small open economy.

One can argue that Proposition 1 is just an application of the Production Efficiency Theorem of Diamond and Mirrlees (1971). Still, *a priori* it is not clear whether the Production Efficiency Theorem applies in the present context where the taxpayer earns quasi-pure ability rent income, Y>0. Notice that Proposition 1 is obtained without requiring Y to be skimmed off by taxation.

4. Second-best efficient policy

In its reduced form, the tax planner's problem is to maximize revenue

$$T = [\frac{1}{\rho}(\omega_1 - w_1)e_1 + \frac{1}{\rho}(\varphi - f)e_\varphi] + [(\omega_2 - w_2)e_2 + (\rho - r)e_\rho]/r \qquad (10)$$

in $x = \varphi, \omega_1, \omega_2, \rho$ subject to the individual budget constraint (6'). In the Appendix it is shown: 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:

$$\frac{\Delta E}{E} = \frac{\Delta L_1 - \Delta E}{L_1 - E} = \frac{\Delta(HL_2)}{HL_2} = \frac{\Delta C_1 - \omega_1\Delta(L_1 - E) + \varphi\Delta E}{C_1 - \omega_1(L_1 - E) + \varphi E} \qquad (11)$$

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

$$\Delta X \equiv \frac{1}{\rho}(\omega_1 - w_1)X_1 + \frac{1}{\rho}(\varphi - f)X_\varphi + \frac{1}{r}(\omega_2 - w_2)X_2 + \frac{\rho - r}{r}X_\rho . \quad (12)$$

According to (12) 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 efficiently chosen. By relying on some simple algebraic manipulations and by making use of

$$\frac{\Delta(HL_2)}{HL_2} = \frac{\Delta L_2}{L_2} + \frac{\Delta H}{H} = \frac{\Delta L_2}{L_2} + \eta \frac{\Delta E}{E}, \quad (13)$$

where the elasticity $\eta \equiv EH'/H$ may well be non-constant in E, (11) can be restated as

$$\frac{\Delta E}{E} = \frac{\Delta L_1}{L_1} = \frac{\Delta C_1}{C_1} \text{ and } \frac{\Delta L_2}{L_2} = (1\text{-}\eta)\frac{\Delta L_1}{L_1}. \quad (11')$$

As differentiation is additive, (11’) could equally and 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 (11’) is shown to imply

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

Hence quantities $C_1, C_2, L_1, E, L_1 - E$ and HL_2 should be reduced in the same proportion relative to the pre-tax position 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-η 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 see that it fails to agree with Ramsey’s Rule of reducing *all* household choices equiproportionately. Only effective labour HL_2 is reduced equiproportionately. As H=H(E) reacts elastically, L_2 is reduced less.

Proposition 2: Second-best efficient policy requires reducing

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

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

$$\frac{\Delta L_2}{L_2} = (1-\eta)\,\frac{\Delta L_1}{L_1} \ . \tag{14}$$

Proposition 2 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 should not expect any interesting relationship to hold in full generality. Still, a remarkably strong result is obtained if the individual learning function is isoelastic, $H(E) = hE^{\eta}$ with $0<\eta<1$. In this special case, the relative reductions $\Delta L_2 / L_2$ and $\Delta E / E$ take on a linear relationship for arbitrary choices of $\varphi, \omega_1, \omega_2, \rho, \pi = \rho(\omega_1 + \varphi)$ and $p = r(w_1 + f)$. This is easily seen when applying the operator Δ to the first-order condition determining the optimal amount of education, $\pi / \omega_2 = H' L_2 = \eta h E^{\eta-1} L_2$:

$$(\eta-1)\frac{\Delta E}{E} + \frac{\Delta L_2}{L_2} = \frac{\Delta(\pi/\omega_2)}{\pi/\omega_2}$$

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

$$= \frac{1}{r}[\frac{w_2}{\omega_2} - \frac{p}{\pi}] \ . \tag{15}$$

Notice that (15) holds for constant η but not necessarily efficient choices of $\varphi, \omega_1, \omega_2$ and ρ while (11’) holds for efficient choices of $\varphi, \omega_1, \omega_2$ and ρ but not necessarily constant η. Comparing (15) with (11’) implies (4).

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

Proposition 3 is the first version of the *Education Efficiency Proposition* derived in this paper. An intuitive explanation is the following. The planner's problem is to set incentives so that two objectives are achieved simultaneously. One objective is to minimize the efficiency loss resulting from distorted choices of utility generating quantities C_1, C_2, L_1 and L_2. The other objective is to minimize losses of quasi-pure ability rent income. In general, these two minimizations are not separable so that the planner has to trade off. Separability is only ensured if (4) holds. If (4) is violated, maximizing the private ability rent in *E* at given L_2 does not necessarily maximize the social ability rent. This is revealed by the following identity:

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

Vice versa, if (4) holds, the last bracketed term on the right-hand side vanishes. This implies that the social ability rent is maximized whenever the private ability rent is maximized by the price-taking taxpayer. Furthermore, because of (15) this maximization need not be traded off against distortions in the choice of L_2. The relative reductions of *E* and L_2 are efficient for arbitrary choices of $\varphi, \omega_1, \omega_2$ and ρ whenever (4) holds. However note that the derivation of (15) requires η being constant so that separability is only obtained for an isoelastic learning function.

Combining Propositions 2 and 3 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 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.

Consider the question of when it is efficient not to distort saving. From Atkinson and Stiglitz (1972) and Sandmo (1974) it is known that it is largely a matter of preferences whether savings should be taxed or not in a Ramsey model of finite periods. This result extends to the present framework:

Proposition 4: If $U(C_1,C_2,L_1,L_2)=U(G(C_1,C_2),L_1,L_2)$ with some homothetic function G, it is efficient not to distort saving and to set $\rho = r$.

The proof is straightforward. Linear homogeneity of G ensures that the ratio of optimal individual consumption does only depend on ρ and on no other policy instrument: $C_1/C_2 = c(\rho)>0$ with $c'<0$. Relying on Proposition 2 and making use of

$$\frac{\Delta C_1}{C_1} - \frac{\Delta C_2}{C_2} = \frac{\Delta[c(\rho)C_2]}{c(\rho)C_2} - \frac{\Delta C_2}{C_2} = \frac{\Delta c}{c} = \frac{\rho - r}{r}\frac{c'}{c}, \tag{16}$$

one obtains $\rho = r$.

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

$$H = E^{\eta},\ U = G(C_1,C_2) - V_1(L_1) - V_2(L_2),\ \nu_i \equiv L_i V_i'' / V_i' \ (i\text{=}1,2), \tag{17}$$

and homothetic G. 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 $Y(\omega_2,\pi,L_2) \equiv \max_E[\omega_2 H(E)L_2 - \pi E]$ in L_2. 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

$$\frac{\tau_2}{\tau_1} = \frac{(1-\eta)\nu_2 - \eta}{\nu_1}. \tag{18}$$

As $\nu_2 > \eta/(1-\eta)$ is to hold by assumption, the numerator on the right-hand side of (18) is positive. For $\eta = 0$, (18) 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 (18) to cope for endogenous education. The effect of education is to reduce τ_2 relative to τ_1. Just note that $(1-\eta)\nu_2 - \eta < \nu_2$.

5. Third-best efficient policies

In this section it is argued that the setting of efficient incentives for education is an objective policy should pursue with higher priority than the setting of efficient incentives for saving. The argument relies on assuming that the set of potential policy instruments $\{\varphi, \omega_1, \omega_2, \rho\}$ is incomplete and on identifying those first-best efficiency conditions which survive under such conditions. It is shown that there are subsets of available instruments for which it is efficient not to distort education even though saving is distorted. This contrasts with all other examined scenarios for which it is efficient to distort saving jointly with education. In particular, no scenario can be identified for which it is efficient not to distort saving but to distort education. In order to ensure an unbiased comparison of education and saving policies, the assumptions $H(E) = hE^{\eta}$ and $U(C_1, C_2, L_1, L_2) = U(G(C_1, C_2), L_1, L_2)$ shall hold throughout. Hence it is second best by assumption neither to distort saving nor to distort education (Propositions 3 and 4). The obvious advantage of making both assumptions and not just one is that it saves one to argue which of the two is empirically more supported.

Proposition 5: Whenever φ and ω_2 are available policy instruments, it is efficient not to distort education even if saving should be distorted for some exogenous reason.

The proof is implicit in the proof of Proposition 3. Just note that the proof of Proposition 3 makes no use of the first-order conditions associated with ω_1 and ρ. The proof of the following proposition is more involved and therefore relegated to the Appendix.

Proposition 6: Whenever either φ or ω_2 is the only non-available policy instrument, it is efficient to distort saving if, and only if, it is efficient to distort education.

An intuitive interpretation for Propositions 5 and 6 is as follows. In the discussion of Proposition 3 it has been argued that the planner has to pursue two objectives simultaneously. One objective is to minimize the efficiency loss resulting from distorted choices of utility generating quantities, while the other objective is to maximize the ability rent. If the learning function is isoelastic and if the set of policy instruments is sufficiently rich, these optimizations are separable. The present section identifies those policy instruments which must be available. The planner must be able to rely on φ and ω_2. While ω_2 allows the planner to target distortions in the choices of utility generating quantities, φ allows her to target the ability rent. If one of the two instruments is lacking, the planner has to trade off the two objectives.

The present analysis is related to studies by Nielsen and Sörensen (1997) and Jacobs and Bovenberg (2007). These authors analyse the merits of dual income taxation. The main result of Nielsen and Sörensen (1997, p. 325) states that labour income should optimally be taxed progressively, $\omega_2 / w_2 < \omega_1 / w_1$, if qualified labour supply is not too elastic and cross substitution of complementarity effects are not too strong. An important qualification of this result is that Nielsen and Sörensen do not model costs of education. Furthermore, they assume that saving is taxed for some exogenous reason. Hence it is as if the two instruments ρ and φ are politically not available. In terms of the present section this means that the progressivity result has at most third-best if not forth-best status. Within the Ramsey framework the instrument φ cannot be substituted by ω_1 without affecting efficiency. Subsidizing the cost of foregone earnings is instrumentally not equivalent to subsidizing the monetary cost of education.

6. The case of heterogeneous taxpayers

In a series of papers Bovenberg and Jacobs (2005) and Jacobs and Bovenberg (2007 and 2008) work out the conditions under which it is optimal (not) to distort education when following the approach of Mirrlees (1971) characterized by asymmetric information. The results derived depend on the availability of nonlinear or only linear policy instruments. Some of the results derived for linear instruments come close to present ones while others do not. The deviations are caused by various differences in the way the planner problem is set up. In some respects, the model of Jacobs and Bovenberg is less general than the present one. E.g. qualified and nonqualified labour supplies are not differentiated and the planner is assumed to have a poll tax at her disposal. Under such assumptions no result like Proposition 2 can be derived. In other respects, the model of Jacobs and Bovenberg is even more general. It is one of heterogeneous taxpayers and differentiated degrees of verifiable learning. *Inter alia* they prove that education should be distorted jointly with saving if education is not verifiable (Jacobs et al., 2007). This strongly reminds one of the part of Proposition 6 referring to the lacking availability of φ. Furthermore, they demonstrate that even a planner trading off efficiency against equity will not compromise on efficiency in education if only education is fully verifiable, the learning function weakly separable and isoelastic (Jacobs at al., 2008). Proposition 7 below confirms and extends this result. The extension lies in showing that two assumptions on which the analysis of Jacobs and Bovenberg is based can be abandoned. One is the availability of a poll tax and the other is the assumption of identical utility functions.

Let $n = 1,..,N$ be the parameter identifying a particular taxpayer. Taxpayers are assumed to differ by preferences and the productivity of learning but not by the elasticity of learning. Hence $u_n = U^n(C_1, C_2, L_1, L_2)$ and $H^n = h_n E^\eta$. 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 (10). In order to model redistribution assume that *n* receives some exogenous income g_n financed out of general tax revenues. The planner then maximizes net aggregate tax revenue subject to the constraints that individual budgets are balanced and that welfare *W* remains constant:

$$\sum_n [T^n - g_n] \text{ in } \varphi, \omega_2, u_n \text{ subject to } g_n = e(\omega_1, \omega_2, \rho, \varphi; u_n) \qquad (\lambda_n)$$

and $W(u_1,..,u_N)$ = constant.

Proposition 7: 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.

The proof is a straightforward extension of the one of Proposition 5. See Appendix.

7. Nonlinear instruments and the Power Law of Learning

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 no 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 the one of 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 $\mathscr{T}$ on qualified labour income and also to subsidize costs of education by some nonlinear scheme S. (ii) Utility functions must be weakly separable in qualified labour and all other individual choice variables. In 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, $\mathscr{T}$'=S'. As a result is not only education undistorted but also saving and nonqualified labour supply. The most direct way of implementing such an optimal taxtransfer system would be the following. (i) Only qualified labour income is taxed. (ii) Taxpayers are allowed to carry forward costs of learning and education and to deduct them against qualified labour income Z_2. Notice that not only foregone earnings should be tax deductible but also all monetary costs of education. See also Trostel (1996 and 1993) who argues in favour of deductions even exceeding one hundred percent.

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, $Z_2 = w_2(n)H(E)L_2$. Jacobs and Bovenberg (2008) find such a specification restrictive and presumably too restrictive to serve as the basis of policy recommendations. However, I find this view too pessimistic given the empirical evidence provided by cognitive psychology. The relevant keyword is the *Power Law of Learning*. 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 by their elasticities than by their level (Anderson, 2005 (1980), Chap. 6; Crossman, 1959). If practice is denoted by *E* and the inverse of the completion time by *H*, this evidence suggests specifying $w_2(n, E_2)$ as $w_2(n)H(E) = w_2 h_n E^{\eta}$. The only drawback is that elasticities may 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 references can be found in the economics literature. A well-known exemption 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.

8. 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 states that Ramsey’s Rule does not apply to qualified labour. Qualified labour supply should be reduced less than nonqualified labour. Only the latter should be reduced in the same proportion as consumption and saving as suggested by Ramsey’s Rule (Proposition 2). No particularly selected utility functions are needed to derive the 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. The familiar Inverse Elasticity Rule is an example which only holds for well selected utility functions. In Section 4 it is shown how this rule has to be adapted if applied to qualified and nonqualified labour. Furthermore, scenarios are considered for which it is efficient not to distort education or saving. In accordance with Atkinson and Stiglitz (1972) and others it is shown to be a matter of taxpayer’s preferences whether saving should be taxed or not (Proposition 4). The Education Efficiency Proposition according to which it is second-best not to distort education is shown to hold if the learning function is isoelastic (Proposition 3). Section 5 analyses third-best policies and provides evidence for the thesis that efficient policy cares more about education than saving. The analysis is third-best in the sense that the set of policy instruments sustaining second-best solutions is incomplete. It is shown that there are scenarios for which it is efficient not to distort education even though saving is distorted. This contrasts with all other examined scenarios for which it is efficient to distort saving jointly with education. In particular, no scenario can be identified for which it is efficient not to distort saving but to distort education. In Section 6 it is finally shown to be efficient not to distort education even if taxpayers are heterogeneous (Proposition 7). This version of the Education Efficiency Proposition generalizes earlier ones by Bovenberg and Jacobs (2005) and Jacobs and Bovenberg (2007 and 2008). 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. (These assumptions do not have to hold in full strength if the Education Efficiency Proposition is proved in the Mirrlees framework. See Jacobs and Bovenberg, 2008, and Anderberg, 2008.)

The theoretical analysis raises the question of whether and to what extent the Education Efficiency Proposition can serve as a guide for education policy. Jacobs and Bovenberg (2008) express scepticism. They find the assumptions needed to prove efficiency in education “restrictive”. In Section 7 a more positive view is proposed. It is argued that cognitive psychology provides impressive evidence for learning functions the elasticity of which does vary neither in the amount of learning nor 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 suggested policy conclusion is that educational choices should at least not be distorted at the intensive margin. Things may be different at the extensive margin. Whether and when it is optimal to intervene with individual choices between competing learning programs cannot however be answered by present study and must be the subject of future research.

9. Appendix

The *proof of* (11) relies on taking partial derivatives of the Lagrange function $T-\lambda e$ with respect to φ, ρ, ω_1, and ω_2:

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

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

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

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), (11′) and (3):

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

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

$$\underset{(11')}{=} \frac{1}{C_2}\frac{\Delta E}{E}\,[\rho\omega_1 L_1+(1-\eta)\omega_2 H L_2-\rho C_1]$$

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

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

Proof of (18): 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 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 (11') yields

$$v_2\frac{\Delta L_2}{L_2} = \frac{\Delta V_2^{'}}{V_2^{'}} = \frac{\Delta\omega_2}{\omega_2} + \eta\frac{\Delta E}{E} \underset{(11')}{=} \frac{1}{r}\frac{\omega_2 - w_2}{\omega_2} + \frac{\eta}{1-\eta}\frac{\Delta L_2}{L_2}$$

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

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 (18).□

Proof of Proposition 6: Assume that ω_2 is the only non-available policy instrument. The planner then maximizes with respect to φ, ω_1, ρ. In (11') only the first two equalities hold:

$$\frac{\Delta E}{E} \;=\; \frac{\Delta L_1}{L_1} \;=\; \frac{\Delta C_1}{C_1}.$$

Additionally use can be made of (16). Writing $\delta = \frac{1}{r}(\frac{w_2}{\omega_2} - \frac{p}{\pi})$, plugging all this into (22) and making use of (15) yields

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

$$\underset{(15),(18)}{=} [\rho C_1 + C_2]\frac{\Delta C_1}{C_1} - \rho\omega_1 L_1\frac{\Delta L_1}{L_1} - \omega_2 HL_2[(1-\eta)\frac{\Delta E}{E} + \delta] \tag{23}$$

$$= [\rho C_1 + C_2 - \rho\omega_1 L_1 - \omega_2 HL_2(1-\eta)]\frac{\Delta E}{E} - \omega_2 HL_2\,\delta$$

$$= -\omega_2 HL_2\,\delta. \tag{24}$$

The latter equality relies on the budget constraint, e=0, and the fact that for an isoelastic learning function and an optimal choice of E the equality $\omega_2 HL_2(1-\eta)$ = $\omega_2 HL_2 - \pi E$ has to hold. When comparing the extreme hand sides of (24) one ends up with $\rho < r \;\Leftrightarrow\; \delta<0 \;\Leftrightarrow\; \frac{w_2}{\omega_2} < \frac{p}{\pi}$.

If φ is the only non-available policy instrument, one has to study the system of equations made up by the first-order conditions

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

and (15). This system is solved for $\Delta L_1, \Delta C_1$ and ΔL_2 as functions of ΔE. One obtains $\frac{\Delta L_1}{L_1} = \frac{\Delta E}{E} + \delta\frac{L_1 - E}{L_1}$, $\frac{\Delta C_1}{C_1} = \frac{\Delta E}{E} + \delta[1 + \varphi\frac{E}{C_1}]$ and $\frac{\Delta L_2}{L_2} = (1-\eta)\frac{\Delta E}{E} + \delta$. Plugging the expressions into (23) yields

$$\frac{\rho - r}{r}\frac{c'}{c}C_2 = [\rho C_1 + C_2]\frac{\Delta C_1}{C_1} - \rho\omega_1 L_1\frac{\Delta L_1}{L_1} - \omega_2 HL_2[(1-\eta)\frac{\Delta E}{E} + \delta]$$

$$= [\rho C_1 + C_2 - \rho\omega_1 L_1 - \omega_2 HL_2(1-\eta)]\frac{\Delta E}{E}$$

$$+ \delta[\rho C_1 + C_2 + \rho\varphi E + \frac{C_2}{C_1}\varphi E - \rho\omega_1(L_1 - E) - \omega_2 HL_2]$$

$$= \delta \frac{C_2}{C_1} \varphi E \,.$$

Hence $\rho < r \;\Leftrightarrow\; \delta > 0 \;\Leftrightarrow\; \frac{w_2}{\omega_2} > \frac{p}{\pi}$ □

The *Proof of Proposition* 7 generalizes the one given above in proving (11). The first-order condition of the planner's maximization with respect to φ is obtained along the lines indicated by (19) and (20):

$$0 \;=\; \rho \sum_n E_n [\frac{\Delta E_n}{E_n} - \lambda_n + \frac{1}{r}]. \qquad (25)$$

The derivation with respect to ω_2 yields:

$$0 \;=\; \sum_n H^n L_{2n} [\frac{\Delta(H^n L_{2n})}{H^n L_{2n}} - \lambda_n + \frac{1}{r}]$$

$$\underset{(13)}{=} \;\; \sum_n H^n L_{2n} [\eta \frac{\Delta E_n}{E_n} + \frac{\Delta L_{2n}}{L_{2n}} - \lambda_n + \frac{1}{r}]$$

$$\underset{(15)}{=} \;\; \sum_n \frac{\pi}{\omega_2 \eta} E_n [\frac{\Delta E_n}{E_n} + \frac{1}{r}(\frac{w_2}{\omega_2} - \frac{p}{\pi}) - \lambda_n + \frac{1}{r}]$$

$$\underset{(25)}{=} \;\; \frac{\pi}{\omega_2 \eta} \frac{1}{r} (\frac{w_2}{\omega_2} - \frac{p}{\pi}) \sum_n E_n$$

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

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, 2007, Human capital and optimal positive taxation of capital income, mimeo: Tinbergen Institute.

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.

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.

Richter, W.F., 2008, Efficient tax policy ranks learning higher than saving, University of Dortmund, 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, 701-6.

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.

CESifo Working Paper Series

for full list see www.cesifo-group.org/wp
(address: Poschingerstr. 5, 81679 Munich, Germany, office@cesifo.de)

2521 Geir B. Asheim and Tapan Mitra, Sustainability and Discounted Utilitarianism in Models of Economic Growth, January 2009

2522 Etienne Farvaque and Gaël Lagadec, Electoral Control when Policies are for Sale, January 2009

2523 Nicholas Barr and Peter Diamond, Reforming Pensions, January 2009

2524 Eric A. Hanushek and Ludger Woessmann, Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation, January 2009

2525 Richard Arnott and Eren Inci, The Stability of Downtown Parking and Traffic Congestion, January 2009

2526 John Whalley, Jun Yu and Shunming Zhang, Trade Retaliation in a Monetary-Trade Model, January 2009

2527 Mathias Hoffmann and Thomas Nitschka, Securitization of Mortgage Debt, Asset Prices and International Risk Sharing, January 2009

2528 Steven Brakman and Harry Garretsen, Trade and Geography: Paul Krugman and the 2008 Nobel Prize in Economics, January 2009

2529 Bas Jacobs, Dirk Schindler and Hongyan Yang, Optimal Taxation of Risky Human Capital, January 2009

2530 Annette Alstadsæter and Erik Fjærli, Neutral Taxation of Shareholder Income? Corporate Responses to an Announced Dividend Tax, January 2009

2531 Bruno S. Frey and Susanne Neckermann, Academics Appreciate Awards – A New Aspect of Incentives in Research, January 2009

2532 Nannette Lindenberg and Frank Westermann, Common Trends and Common Cycles among Interest Rates of the G7-Countries, January 2009

2533 Erkki Koskela and Jan König, The Role of Profit Sharing in a Dual Labour Market with Flexible Outsourcing, January 2009

2534 Tomasz Michalak, Jacob Engwerda and Joseph Plasmans, Strategic Interactions between Fiscal and Monetary Authorities in a Multi-Country New-Keynesian Model of a Monetary Union, January 2009

2535 Michael Overesch and Johannes Rincke, What Drives Corporate Tax Rates Down? A Reassessment of Globalization, Tax Competition, and Dynamic Adjustment to Shocks, February 2009

2536 Xenia Matschke and Anja Schöttner, Antidumping as Strategic Trade Policy Under Asymmetric Information, February 2009

2537 John Whalley, Weimin Zhou and Xiaopeng An, Chinese Experience with Global 3G Standard-Setting, February 2009

2538 Claus Thustrup Kreiner and Nicolaj Verdelin, Optimal Provision of Public Goods: A Synthesis, February 2009

2539 Jerome L. Stein, Application of Stochastic Optimal Control to Financial Market Debt Crises, February 2009

2540 Lars P. Feld and Jost H. Heckemeyer, FDI and Taxation: A Meta-Study, February 2009

2541 Philipp C. Bauer and Regina T. Riphahn, Age at School Entry and Intergenerational Educational Mobility, February 2009

2542 Thomas Eichner and Rüdiger Pethig, Carbon Leakage, the Green Paradox and Perfect Future Markets, February 2009

2543 M. Hashem Pesaran, Andreas Pick and Allan Timmermann, Variable Selection and Inference for Multi-period Forecasting Problems, February 2009

2544 Mathias Hoffmann and Iryna Shcherbakova, Consumption Risk Sharing over the Business Cycle: the Role of Small Firms' Access to Credit Markets, February 2009

2545 John Beirne, Guglielmo Maria Caporale, Marianne Schulze-Ghattas and Nicola Spagnolo, Volatility Spillovers and Contagion from Mature to Emerging Stock Markets, February 2009

2546 Ali Bayar and Bram Smeets, Economic and Political Determinants of Budget Deficits in the European Union: A Dynamic Random Coefficient Approach, February 2009

2547 Jan K. Brueckner and Anming Zhang, Airline Emission Charges: Effects on Airfares, Service Quality, and Aircraft Design, February 2009

2548 Dolores Messer and Stefan C. Wolter, Money Matters – Evidence from a Large-Scale Randomized Field Experiment with Vouchers for Adult Training, February 2009

2549 Johannes Rincke and Christian Traxler, Deterrence through Word of Mouth, February 2009

2550 Gabriella Legrenzi, Asymmetric and Non-Linear Adjustments in Local Fiscal Policy, February 2009

2551 Bruno S. Frey, David A. Savage and Benno Torgler, Surviving the Titanic Disaster: Economic, Natural and Social Determinants, February 2009

2552 Per Engström, Patrik Hesselius and Bertil Holmlund, Vacancy Referrals, Job Search, and the Duration of Unemployment: A Randomized Experiment, February 2009

2553 Giorgio Bellettini, Carlotta Berti Ceroni and Giovanni Prarolo, Political Persistence, Connections and Economic Growth, February 2009

2554 Steinar Holden and Fredrik Wulfsberg, Wage Rigidity, Institutions, and Inflation, February 2009

2555 Alexander Haupt and Tim Krieger, The Role of Mobility in Tax and Subsidy Competition, February 2009

2556 Harald Badinger and Peter Egger, Estimation of Higher-Order Spatial Autoregressive Panel Data Error Component Models, February 2009

2557 Christian Keuschnigg, Corporate Taxation and the Welfare State, February 2009

2558 Marcel Gérard, Hubert Jayet and Sonia Paty, Tax Interactions among Belgian Municipalities: Does Language Matter?, February 2009

2559 António Afonso and Christophe Rault, Budgetary and External Imbalances Relationship: A Panel Data Diagnostic, February 2009

2560 Stefan Krasa and Mattias Polborn, Political Competition between Differentiated Candidates, February 2009

2561 Carsten Hefeker, Taxation, Corruption and the Exchange Rate Regime, February 2009

2562 Jiahua Che and Gerald Willmann, The Economics of a Multilateral Investment Agreement, February 2009

2563 Scott Alan Carson, Demographic, Residential, and Socioeconomic Effects on the Distribution of 19th Century US White Statures, February 2009

2564 Philipp Harms, Oliver Lorz and Dieter Urban, Offshoring along the Production Chain, February 2009

2565 Patricia Apps, Ngo Van Long and Ray Rees, Optimal Piecewise Linear Income Taxation, February 2009

2566 John Whalley and Shunming Zhang, On the Arbitrariness of Consumption, February 2009

2567 Marie-Louise Leroux, Endogenous Differential Mortality, Non-Contractible Effort and Non Linear Taxation, March 2009

2568 Joanna Bęza-Bojanowska and Ronald MacDonald, The Behavioural Zloty/Euro Equilibrium Exchange Rate, March 2009

2569 Bart Cockx and Matteo Picchio, Are Short-Lived Jobs Stepping Stones to Long-Lasting Jobs?, March 2009

2570 David Card, Jochen Kluve and Andrea Weber, Active Labor Market Policy Evaluations: A Meta-analysis, March 2009

2571 Frederick van der Ploeg and Anthony J. Venables, Harnessing Windfall Revenues: Optimal Policies for Resource-Rich Developing Economies, March 2009

2572 Ondřej Schneider, Reforming Pensions in Europe: Economic Fundamentals and Political Factors, March 2009

2573 Jo Thori Lind, Karl Ove Moene and Fredrik Willumsen, Opium for the Masses? Conflict-Induced Narcotics Production in Afghanistan, March 2009

2574 Silvia Marchesi, Laura Sabani and Axel Dreher, Agency and Communication in IMF Conditional Lending: Theory and Empirical Evidence, March 2009

2575 Carlo Altavilla and Matteo Ciccarelli, The Effects of Monetary Policy on Unemployment Dynamics under Model Uncertainty - Evidence from the US and the Euro Area, March 2009

2576 Falko Fecht, Kjell G. Nyborg and Jörg Rocholl, The Price of Liquidity: Bank Characteristics and Market Conditions, March 2009

2577 Giorgio Bellettini and Filippo Taddei, Real Estate Prices and the Importance of Bequest Taxation, March 2009

2578 Annette Bergemann and Regina T. Riphahn, Female Labor Supply and Parental Leave Benefits – The Causal Effect of Paying Higher Transfers for a Shorter Period of Time, March 2009

2579 Thomas Eichner and Rüdiger Pethig, EU-Type Carbon Emissions Trade and the Distributional Impact of Overlapping Emissions Taxes, March 2009

2580 Antonios Antypas, Guglielmo Maria Caporale, Nikolaos Kourogenis and Nikitas Pittis, Selectivity, Market Timing and the Morningstar Star-Rating System, March 2009

2581 António Afonso and Christophe Rault, Bootstrap Panel Granger-Causality between Government Budget and External Deficits for the EU, March 2009

2582 Bernd Süssmuth, Malte Heyne and Wolfgang Maennig, Induced Civic Pride and Integration, March 2009

2583 Martin Peitz and Markus Reisinger, Indirect Taxation in Vertical Oligopoly, March 2009

2584 Petra M. Geraats, Trends in Monetary Policy Transparency, March 2009

2585 Johannes Abeler, Armin Falk, Lorenz Götte and David Huffman, Reference Points and Effort Provision, March 2009

2586 Wolfram F. Richter, Taxing Education in Ramsey's Tradition, March 2009