Education: Optimal choice and efficient policy

Wolfram F. Richter
TU Dortmund University, IWH Halle, CESifo Munich and IZA Bonn

Kerstin Schneider
WIB University of Wuppertal and CESifo Munich

Abstract. This paper argues that it suffices to assume distortionary wage taxation to prove the efficiency of effective subsidization of education. The paper does not rely on considerations of equity and market failure to justify subsidies. Instead, the optimal subsidy reduces the social cost of distortive wage taxation. The theoretical approach assumes a Mincer-type earnings function, analyzes corner solutions of optimal schooling choice and derives the result of efficient subsidization in a Ramsey-type framework. Second-best policy is confronted with empirical evidence from OECD countries. The majority of countries are shown to subsidize tertiary education in effective terms.

Résumé. Éducation : choix optimal et politique efficace. Afin de démontrer l'efficacité d'un subventionnement effectif de l'éducation, nous affirmons qu'il suffit de supposer une distorsion des salaires. Pour justifier ces subventions, nous écartons toute considération d'équité ou de déficience du marché. À la place, nous affirmons qu'un subventionnement optimal permet de réduire le coût social lié à la distorsion de la fiscalité sur les salaires. L'approche théorique s'appuie sur une fonction de gains de Mincer, évalue les solutions d'angle relativement aux choix d'études optimaux puis extrapole le résultat d'un subventionnement efficace dans un modèle de Ramsey. La politique de second choix est comparée aux observations empiriques réalisées au sein pays de l'OCDE. Il apparaît que la plupart des pays subventionnent l'enseignement supérieur en termes effectifs.

JEL classification: H21, I28, J24

1. Introduction

GROWTH AND WELFARE in the knowledge society depends on countries' investments in human capital, and because the private expected rates of return to education are estimated to be high, rational individuals should invest in education. But if investment in education is in the individual's self-

---

Corresponding author: Wolfram F. Richter, Wolfram.Richter@tu-dortmund.de

Helpful comments of editor Arthur Sweetman and anonymous referees are gratefully acknowledged.

Canadian Journal of Economics / *Revue canadienne d'économique* 2021 54(2)
May 2021. Printed in Canada / *Mai 2021. Imprimé au Canada*

ISSN: 0008-4085 / 20 / pp. 840–863 / DOI: 10.1111/caje.12518

interest, is there a role for efficiency enhancing government intervention in the absence of market failures? And if so, should higher education be taxed or subsidized? Two strands of literature deal with these questions. One is positive theoretic in spirit, with primary focus on estimating the returns to schooling. The other has grown out of optimal tax theory. It is normative theoretic and characterizes optimal education policy. Although both deal with related issues, it is surprising that only a small number of contributions have tried to merge these two strands of literature. In essence, that is what this paper performs. The main contribution of the paper is the development of an optimal tax model in the tradition of Ramsey that integrates the individual's schooling decision with a convex earnings function. We then test the implications using data from OECD countries.

The traditional approach to modelling the individual schooling decision builds on Mincer (1958) and Becker (1964). It relies on the assumption that individuals maximize the present value of lifetime earnings. Although appealing, the idea that schooling decision results from strict income maximizing behaviour is challenged by evidence of significant non-pecuniary returns and costs of education. In fact, Heckman et al. (2006) describe the evidence against strict income maximization as "overwhelming." Summarizing the literature on the non-pecuniary returns, Oreopoulos et al. (2011, p. 180) conclude that non-pecuniary returns "are both real and important." As to the non-pecuniary costs of education, Heckman et al. (2006, p. 436) suggest that psychic costs "play a very important role." However, the black-box character of non-pecuniary returns and costs is not satisfactory (Heckman et al. 2006, p. 436). Attempts to explain schooling decision with vague notions of psychic costs is little more than acknowledging the fact that schooling decision is not well understood.

The present paper develops an alternative approach based on the following: (i) schooling decision based on utility maximization rather than pure income maximization, (ii) a Mincer-type earnings function and (iii) an analysis of efficient educational policy as opposed to the attempt to estimate the returns to schooling. None of these components are novel in their own right; however, our contribution, as we would like to convince the reader, comes from combining all three of them.

Utility maximization has a natural appeal as the standard assumption in the neoclassical paradigm of individual behaviour and it also serves as a basis for the analysis of allocational efficiency. Initially, utility and income maximization appear to be concepts with equivalent behavioural implications; hence, one does not expect to get new insights when replacing one with the other. However, utility maximization and income maximization do have different implications when the earnings function is convex and not concave.

Recall two robust results of the empirical literature: First, there is strong empirical evidence for (locally) convex earnings functions and second, estimates of the (Mincer) coefficient of years of schooling notoriously exceed standard real discount rates raising doubts about the rationality of schooling decision (Card 1999). However, this puzzle can be resolved once the decision on the optimal amount of schooling is modelled in a utility maximization framework with rational individuals and earnings functions that are convex at the optimum. As will be argued below utility maximization implies that the return to schooling is equal to the cost of forgone leisure, which systematically exceeds the cost of foregone income. Thus, the true marginal internal rate of return to schooling is simply overestimated when measuring the cost of schooling with the observable cost of foregone earnings instead of the unobservable cost of foregone leisure.

The analysis of optimal schooling choice with convex earnings functions is only the first part of the analysis. The paper's ultimate objective is to characterize efficient education policy. To do so, however, one has to deal with corner solutions that result when maximizing multi-period utility subject to a convex earnings function. The literature surveyed in section 2 is not faced with the difficulty of corner solutions by combining either income maximization with convex earnings functions or utility maximization with concave earnings functions. Here, we combine the empirically relevant case of (locally) convex earnings functions with the methodologically convincing assumption of utility maximization.

In this paper, efficient education policy is derived in Ramsey's tradition. As it turns out, the elasticity of the earnings function and the sign of its derivative are of pivotal relevance. It is shown that distortive wage taxation entails subsidizing education in effective terms if the earnings function displays increasing elasticity in the amount of schooling. (Note that any function that is log-linear in schooling like the standard Mincer earnings function will feature convexity as well as increasing elasticity.) Effective subsidization means that the resulting quantity of schooling exceeds the first-best level. Thus, the government's need to raise revenue through a wage tax already renders an effective subsidy for education optimal. Other distortions, like imperfect capital markets, also justify subsidizing tertiary education, but only up to the first-best level and not beyond.

In the paper's empirical section, optimal policy is confronted with evidence from a panel of OECD countries. It is shown that education policies in OECD countries indeed tend towards effective subsidization of tertiary education. There is some first evidence that in some countries the extent of effective subsidization goes beyond the second-best optimum.

The paper is organized as follows. Section 2 briefly reviews the related literature. Section 3 defines the concept of a Mincer-type earnings function on which the theory of the paper is based. Section 4 sets up a standard model of a representative individual who invests in education by maximizing lifetime utility. Section 5 proves that optimal Ramsey policy requires subsidizing education in effective terms. Section 6 confronts second-best policy with empirical evidence from a sample of OECD countries. Section 7 concludes.

2. Related literature

This paper combines two strands of literature. The older strand has emerged from labour and education economics. It was initiated by Mincer (1958) and Becker (1964) and is positive theoretic in spirit. The focus is on schooling decision and earnings determination. The other strand has grown out of the public and the macroeconomics literature. It is normative theoretic and is the starting point for the analysis of the optimal taxation of education. Examples are Bovenberg et al. (2005), Anderberg (2009), Richter (2009), Jacobs et al. (2011) and Krueger et al. (2013).[1] A major shortcoming of this literature is that it assumes concave earnings functions despite empirical evidence from the literature on labour economics reinforcing the convexity of the earnings function. In fact, there is even evidence for growing convexity of the earnings function over time (Lemieux 2006). In order to reconcile the two strands of literature, the normative analysis has to be extended to be applicable for convex earnings functions. This extension is not obvious and will be done in sections 4 and 5 below.

The literature, following Becker and Mincer, estimates earnings functions with schooling being modelled in continuous time. The bottom line of this literature, well surveyed by Card (1999), is that the rate at which earnings grow in years of schooling tends to exceed any realistic real discount rate. This raises a puzzling question: why don't individuals continue schooling despite the high returns?

More recent contributions in the tradition of Roy (1951) and Willis et al. (1979) model schooling decision as a problem of self-selection. In line with the theory of comparative advantage, the individual is assumed to make a discrete choice between continuing or discontinuing schooling.

However, the estimated marginal internal rates of return to schooling still substantially exceed the level of real interest rates (Heckman et al. 2006, Heckman et al. 2008). One possible, and often suggested, explanation refers to liquidity constraints, in particular for marginal students (Zimmerman 2014). However, even though public concerns about credit constraints are strong, the impact of the latter on tertiary education is estimated to be relatively weak (Carneiro et al. 2002). All this has led Heckman et al. (2008) to challenge the assumption that individuals maximize income only when making schooling decisions. They suggest accounting for heterogeneity and including psychic costs in the analysis. Compared with low-ability individuals, more-able individuals are argued to have lower psychic costs of attending college.

A seminal paper by Carneiro et al. (2011) presents returns to education, explicitly accounting for individual observed and unobserved heterogeneity as well as sorting issues. The average treatment effect is lower than the

---

1 There are also intermediate cases such as Wallenius (2011). Her paper is positive theoretic but assumes a Ben-Porath type technology defining skill as a concave function of training.

treatment effect on the treated but substantially higher than the treatment effect on the untreated. Interestingly, the effect on the untreated is below a typically assumed discount rate. Carneiro et al. (2011) estimate the distribution of the marginal treatment effects and a marginal policy relevant treatment effect resulting from a small change in education policy.

However, the focus of the empirical literature is not efficient education policy but the estimation of returns to tertiary education. From the perspective of the present paper, the acknowledged importance of unobservables in explaining the decision to attend college is a key feature because the "unobserved component of the desire to go to college" (Carneiro et al. 2011, p. 2758) suggests that individuals in fact maximize utility rather than income.

Besides the labour economics literature, there is research on schooling decision and education policy in the fields of public economics and the macroeconomics. And surprisingly, there is hardly any cross-acknowledgment between the two fields of literature. A notable exception is a paper by Findeisen et al. (2015). The authors calibrate a model combining optimal nonlinear income taxation in the tradition of Mirrlees (1971) with discrete schooling decision in the tradition of Roy (1951), Willis et al. (1979), Heckman et al. (2006) and others.

Findeisen et al. (2015) follow Mirrlees (1971) and Bovenberg et al. (2005) in allowing for individual heterogeneity. Their model incorporates idiosyncratic risk and borrowing constraints as well as multidimensional heterogeneity. The downside of this complexity is simplicity in modelling details. For example, individual preferences are assumed to be quasi linear. Furthermore, psychic costs, which are not well understood, are pivotal for explaining schooling decision.

In the following, we propose a model that builds on *arbitrary* utility functions and does not rely on unspecified psychic costs. This level of generality comes at the cost of neglecting heterogeneity. However, we argue that disregarding individual heterogeneity is appropriate when analyzing policy issues. After all, tax and education policy is not designed for individuals or small groups characterized by distinct individual criteria. Tax and education policy must set efficient incentives for average individuals.

3. Mincer-type earnings function

In the following, education and schooling are used synonymously. Earnings per unit of time, $G = G(E)$, increase in schooling, E, and are determined by demand and supply in the labour market. Individuals consider the earnings function in the relevant range as given, when deciding on the amount of schooling. Ignoring the effects of work experience on earnings, the *standard Mincer earnings function* is log-linear,

$$\ln G(E) = a + mE, \tag{1}$$

where the parameters a and m are positive constants. A log-linear function has two formal properties: it is (i) convex, $G''(E)>0$, and (ii) of increasing elasticity. More precisely, the elasticity $\gamma(E) \equiv EG'/G$ is increasing in E. In fact, the elasticity of the standard Mincer earnings function is $\gamma(E) = mE$ and increases proportionally in E. Note that convexity and increasing elasticity are mathematically independent properties: functions can be convex and have a decreasing elasticity and functions can have an increasing elasticity but are concave. In this paper, the earnings functions are not restricted to log-linear functions. However, we assume earnings functions that feature convexity and increasing elasticity in the relevant domain of education. Such functions are called *Mincer-type earnings functions.* As will be shown in the following, individuals end up in corner solutions when they maximize lifetime utility subject to a convex earnings function. This means that they invest in either education or supply labour but never do both in the same period of life. The consequence of increasing elasticity is that it is second-best efficient to subsidize education if labour supply is distorted by wage taxes.

4. Household behaviour

In the following, we introduce the Mincer-type earnings function in a standard model of household behaviour that is common in the public economics and the macroeconomic literature. The focus is on a representative individual living for two periods and deriving increasing utility, U, from consumption, C_i, and decreasing utility from non-leisure time, $L_i \leq \overline{L}_i$, in periods $I = 1{,}2$. $\overline{L}_i$ is the length of period i. The function $U = U(C_1, C_2, L_1, L_2)$ is quasi-concave. Non-leisure in period 2, L_2, is second-period labour supply, whereas, in period 1, $L_1 - E$ is labour supply with E being time spent on education. First-period labour earns a constant wage rate after tax, ω_1; the return to second-period labour, however, depends on the amount of education. Workers get paid $\omega_2 G(E)$ per unit of time, where G is a Mincer-type earnings function and ω_2 accounts for wage taxation. I.e. ω_2 equals one minus the wage tax of period 2. Given a positive choice of education, $E > 0$, second-period labour is interpreted as qualified labour. Likewise, the quantities $L_1 - E$ and L_1 are interpreted as non-qualified labour and non-qualified non-leisure, respectively. Education causes opportunity costs in the form of foregone earnings and costs of tuition. Both costs are assumed to be linear in time spent on education. The cost of foregone earnings is modelled by $\omega_1 E$, and the cost of tuition is φE. The share of first-period income that is not spent on education or on consumption is first-period savings:

$$S = \omega_1(L_1 - E) - \varphi E - C_1 = \omega_1 L_1 - (\varphi + \omega_1)E - C_1. \tag{2}$$

By way of normalization, the price of consumption is set equal to one. The gross rate of return to saving is denoted by ρ, and we assume perfect capital markets. In particular, there are no credit constraints; hence, negative savings are no problem. The only inefficiency comes from taxation.

All second-period income is spent on consumption:

$$C_2 = \rho S + \omega_2 G(E) L_2. \tag{3}$$

Substituting for S in equations (2) and (3) yields the lifetime budget constraint:

$$C_1 + C_2/\rho = \omega_1 L_1 + \omega_2 G(E) L_2/\rho - (\varphi + \omega_1)E. \tag{4}$$

Maximizing utility $U(C_1, C_2, L_1, L_2)$ in $C_1, C_2, L_1, L_2, E \geq 0$ subject to constraint (4), $\overline{L}_1 \geq L_1 \geq E$, and $\overline{L}_2 \geq L_2$ requires that net income, Y, is maximized in E holding other variables constant,

$$\begin{aligned} Y(L_1, L_2) &= Y(L_1, L_2; \omega_2/\rho(\varphi + \omega_1)) \\ &\equiv \max_{0 \leq E \leq L_1} [\omega_2 G(E) L_2/\rho - (\varphi + \omega_1)E]. \end{aligned} \tag{5}$$

Equation (5) looks like a discrete version of income maximization à la Mincer and Becker. Note, however, that equation (5) assumes linear costs of education, while the standard Mincer schooling model implicitly assumes increasing costs. This has implications for the characterization of optimal behaviour and needs some careful analysis.

When maximizing equation (5), we can conceive three scenarios. In the first one, it is optimal for the taxpayer to receive no education, i.e., $E = 0$. This is the case whenever the incentive to invest in education is too weak, for instance, because the wage premium is low or the tax on qualified labour is high. In the second scenario, maximizing the net income of education has an *interior solution* with $E \in (0, L_1)$. Obviously, this can happen only if the earnings function is weakly concave, $G''(E) \leq 0$. Concavity of the earnings function, however, has been excluded by assumption. Therefore, with convexity, the *upper corner solution* with $E = L_1$ will be optimal whenever the individual decides to invest in education, $E > 0$. Thus, utility maximizing individuals spend all non-leisure time on either unqualified labour supply or education.

Proposition 1. *With a Mincer-type earnings function and utility maximization, non-leisure time is optimally spent either on working or on education but not on both in the same period.*

Maximizing utility at $E > 0$ implies equating the return to education, $\omega_2 G'(E) L_2/\rho$, with the opportunity cost of education, $\varphi + MRS^1$ where $MRS^1 \equiv -U_{L_1}/U_{C_1}$ is the *cost of foregone leisure*. To put it differently, the *private marginal internal rate of return to education*, IRR_{priv}, and the gross rate of interest have to be equated,

$$IRR_{priv} \equiv \frac{\omega_2 G'(E) L_2}{\varphi + MRS^1} = \rho. \tag{6}$$

This optimality condition is a focal one in the Mincer literature. The pivotal difference to the present approach comes from interpreting MRS^1. The standard Mincer model builds on income maximization. If the choice of education is to be explained by income maximization, this requires that all returns and costs can be expressed in monetary units and that the cost of foregone leisure, MRS^1, can be equated with the *cost of foregone earnings*, ω_1. It has, however, been shown that with utility maximizing individuals this equality holds only when the earnings function is weakly concave. As convexity, the empirically relevant case, has been assumed, utility is maximized at an upper corner solution and the cost of foregone leisure necessarily exceeds the cost of foregone earnings, $MRS^1 > \omega_1$. Thus, IRR_{priv} is *systematically* overestimated when the cost of education is proxied by the cost of foregone earnings rather than by the cost of foregone leisure.

Maximizing the net income of education, Y, generates increasing returns. This is hardly surprising if the earnings function is convex. Note, however, that increasing returns would also result with concave earnings functions and interior solutions. In this case, net income, Y, is convex in L_2:

$$\frac{d^2Y}{dL_2^2} = \frac{\omega_2}{\rho} G' \frac{dE}{dL_2} = -\frac{\omega_2}{\rho} \frac{(G')^2}{G'' L_2} > 0. \tag{7.a}$$

With convex earnings functions, we always get upper-corner solutions, and Y is convex in L_1:

$$\frac{d^2Y}{dL_1^2} = \frac{\omega_2}{\rho} G'' L_2 > 0. \tag{7.b}$$

The convexity of the net income function, Y, has implications for the individual's optimization. Just assuming quasi-concavity of the utility function is clearly not sufficient to ensure that the individual's optimization is well behaved. The second-order conditions are not necessarily satisfied and interior solutions of $L_i \in [0, \overline{L}_i]$ may fail to exist. Still, the following analysis looks at only first-order conditions. The implicit assumption is, first, that the individual discards all solutions of the first-order conditions that fail to be globally optimal and, second, that Inada-type conditions hold. The latter implies that marginal disutility of non-leisure tends to infinity when L_i approaches the upper bound, $\overline{L}_i$, and that marginal disutility of non-leisure tends to zero when L_i approaches zero.

One might conjecture that corner solutions are an artefact of the two-period model and not in line with empirical evidence. This is, however, not correct. In a multi-period version, the individual has to decide over how many periods to be educated. Earnings in period $i+1$ are a function of earlier education, $G = G(E_1 + .. + E_i)$. In this model, it is still optimal to spend non-leisure time either on education or on working in each period, whenever G is convex. If neither the cost of tuition nor the wage rate decrease in present values over time, $\omega_{i-1} \leq \omega_i/\rho$ and $\varphi_{i-1} \leq \varphi_i/\rho$, there will be a period $\overline{i}$ after which individuals switch from education to work. Thus individuals find it optimal to be educated in all earlier periods, $i < \overline{i}$, and to work in all later periods, $i \geq \overline{i}$. Allowing for heterogeneous individuals, the cut-off period $\overline{i}$ depends on the marginal disutility of non-leisure and differs between individuals, which is in line with the observed variation in time spent on education (the proof is made available as online appendix A).

5. Second-best policy

We now turn to optimal policy design in the two-period model, keeping in mind that the results apply equally to a multi-period setting. The government needs to raise revenue to which end there are four possible linear tax instruments, each of which distorts the individual's decision. Taxes can be levied on labour income in the first and second periods, on the cost of tuition and on the returns to savings. They are modelled implicitly as the difference between prices before and after taxes. The prices after taxes and subsidies are endogenous and denoted by $\omega_1, \omega_2, \varphi, \rho$. The prices before taxes and subsidies are exogenous and denoted by w_1, w_2, f, r.[2] The tax on labour income in period $i=1$, and 2 is modelled by $w_i - \omega_i$, the tax on capital income by $r-\rho$ and the tax on the cost of tuition by $\varphi - f$. It goes without saying that each tax can be negative, i.e., a subsidy. Government's net revenue amounts to

$$T \equiv (w_1 - \omega_1)(L_1 - E) + (\varphi - f)E + [(w_2 - \omega_2)G(E)L_2 + (r-\rho)S]/r. \qquad (8)$$

In order to characterize second-best tax policy, it is convenient to work with the taxpayer's expenditure function, which is defined by

$$e(\omega_1, \omega_2, \varphi, \rho; u) \equiv \min[\rho C_1 + C_2 + \rho(\varphi + \omega_1)E - \rho\omega_1 L_1 - \omega_2 G(E)L_2]. \qquad (9)$$

in C_1, C_2, L_1, L_2, E subject to $U(C_1, C_2, L_1, L_2) \geq u$ and $L_1 \geq E$. Assume that the expenditure function is twice differentiable. Relying on Hotelling's lemma yields the identities $e_{\omega_1} = -\rho(L_1 - E)$, $e_{\omega_2} = -G(E)L_2$, $e_\varphi = \rho E$ and $e_\rho = C_1 + \varphi E = -S$, where the variables L_i, E, S and C_i have to be interpreted as Hicksian supply and demand functions to be evaluated at $\omega_1, \omega_2, \varphi, \rho$ and u. Note that the expenditure function is independent of ω_1, hence $e_{\omega_1} \equiv 0$, when all non-leisure time of period 1 is spent on education, $L_1 = E$, i.e., the individual is at the upper corner solution. Using these definitions, equation (8) can be written as

$$T = \frac{1}{\rho}(\omega_1 - w_1)\, e_{\omega_1} + \frac{1}{\rho}(\varphi - f)\, e_\varphi + \frac{1}{r}\left[(\omega_2 - w_2)\, e_{\omega_2} + (\rho - r)\, e_\rho\right]. \qquad (10)$$

2 The function $G(E)$ has been introduced as earnings, suggesting $w_2 = 1$. If one chose instead to interpret education as a labour augmenting activity and $G(E)L_2$ as *effective qualified labour*, w_2 equalled the latter's marginal productivity. It is a straightforward exercise to endogenize the prices before taxes and subsidies in this case. However, endogenization does not produce interesting new insights. Assuming no pure profits in the private sector so that the production efficiency theorem applies, endogenous prices leave the structure of efficient education policy unchanged.

The planner's objective is to maximize revenue T in $\omega_1,\omega_2,\varphi,\rho$ subject to the taxpayer's budget constraint, $e(\omega_1,\omega_2,\varphi,\rho;u)=0$. In the planner's optimum, all Hicksian demands and supplies, E, C_1, C_2, L_1 and GL_2, appearing in the taxpayer's budget constraint have to be reduced by the same proportion. This requirement is conveniently expressed by the use of hat notation,

$$\widehat{E}=\widehat{L_1}=\widehat{GL_2}=\widehat{C}_1. \tag{11}$$

A hat on a function $X=X(\omega_1,\omega_2,\varphi,\rho;u)$ denotes a relative change, $\widehat{X}\equiv\Delta X/X$, with the total differential operator Δ defined by

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

According to equation (12), ΔX equals the weighted sum of the partial derivatives of X with the weights given by the tax wedges. The efficiency conditions (11) are derived in Richter (2009) for concave earnings functions and they are shown to imply

$$\widehat{C_2}=\widehat{E},\ \widehat{L_2}=(1-\gamma)\widehat{L_1}\text{ and }\widehat{G}=\gamma\widehat{E}, \tag{13}$$

where γ is the elasticity of the earnings function. It is, however, possible to extend the derivations to convex earnings functions, i.e., equations (11) and (13) hold for both concave and convex earnings functions (the proof is made available as online appendix B).

The equi-proportionate reduction of demands and supplies is clearly in line with Ramsey's (1927) characterization of efficient taxation. The less standard result concerns the change in qualified labour supply, L_2. Efficient taxation requires the relative reduction in qualified labour to be smaller than the relative reduction in non-qualified labour. The factor is $1-\gamma$; hence, the tax induced reduction in qualified labour decreases in γ. Thus, the more elastic the individual earnings function, the smaller the optimal reduction in qualified labour relative to non-qualified labour. While this result is quite intuitive, it is clearly in contrast to Ramsey's rule of reducing all household choices equi-proportionately. In the model with endogenous education, *effective* qualified labour, GL_2, is reduced equi-proportionately. Qualified labour, however, should be reduced less than proportionately, as the elasticity of the earnings function γ is positive. For earnings functions with elasticity γ greater than one, L_2 should even increase (cf. equation (13)).

The optimal choice of education is characterized in equation (6). It states the equality of the private marginal internal rate of return to education and the private discount rate. This condition is equivalent to the condition that the *marginal return to education* equals the (*effective*) *marginal cost of education*,

$$MR\equiv G'(E)L_2=\rho\big(\varphi+MRS^1\big)/\omega_2\equiv MC. \tag{14}$$

Applying hat calculus to the left-hand side of equation (14) yields

$$\widehat{MR} = \frac{\Delta MR}{MR} = \frac{\Delta(G'L_2)}{G'L_2} = \frac{\Delta G'}{G'} + \frac{\Delta L_2}{L_2} = \left[\frac{EG''}{G'} + 1 - \gamma\right]\hat{E} = \gamma_\gamma \widehat{E}, \tag{15}$$

where $\gamma_\gamma \equiv E\gamma'/\gamma$ denotes the second-order elasticity of the earnings function. This second-order elasticity is necessarily positive, as the elasticity of the earnings function is assumed to be increasing in E. $\gamma_\gamma$equals one if the earnings function is log-linear (the Mincerian case). As $\widehat{E}$ is negative, given the need to raise positive tax revenue, it follows from equation (15) that the efficient change in the marginal return to education, $\widehat{MR}$, is negative as well. Because MR equals MC, the efficient change in the marginal cost of education has to be negative as well. Applying hat calculus to the right-hand side of equation (14) yields

$$\begin{aligned} 0 > \widehat{MC} &= \frac{\Delta MC}{MC} = \frac{\Delta(\rho\varphi) + \Delta\left(\rho MRS^1\right)}{\rho\left(\varphi + MRS^1\right)} - \frac{\Delta\omega_2}{\omega_2} \\ &= \frac{(\varphi - f) + (\rho - r)\left(\varphi + MRS^1\right)/r + \rho\Delta MRS^1}{\rho\left(\varphi + MRS^1\right)} - \frac{\omega_2 - w_2}{r\omega_2} \\ &= \frac{w_2/r}{\omega_2} - \frac{f + MRS^1 - \rho\Delta MRS^1}{\rho\left(\varphi + MRS^1\right)}. \end{aligned} \tag{16}$$

If $MRS^1 = \omega_1$ holds, this implies $\Delta MRS^1 = \Delta\omega_1 = (\omega_1 - w_1)/\rho$ and $MRS^1 - \rho\Delta MRS^1 = w_1$. However, as argued above, MRS^1 equals ω_1 only if the earnings function is concave. If the earnings function is convex, it is nevertheless suggestive to write $MRS^1 \equiv \omega_1^s$ and $MRS^1 - \rho\Delta MRS^1 \equiv w_1^s$ and to interpret ω_1^s and w_1^s as the respective *private* and *social shadow costs of foregone non-qualified leisure*. Hence, equation (16) can be restated as

$$0 > \rho\widehat{MC} = \frac{w_2/r}{\omega_2/\rho} - \frac{f + w_1^s}{\varphi + \omega_1^s} = \frac{w_2 G' L_2/r - \left(f + w_1^s\right)}{\varphi + \omega_1^s} \equiv \Delta_E, \tag{17}$$

where Δ_E is interpreted as the *effective wedge on education*, and equation (6) has been used. The inequality in equation (17) is equivalent to $w_2 G' L_2/r < f + w_1^s$.

In the following, we refer to *education* as (*effectively*) *subsidized* when the effective wedge on education is negative. A negative wedge requires a level of education for which the (*effective*) *marginal social cost*, $r\left(f + w_1^s\right)/w_2$, exceeds the *marginal social return*, $G'L_2$. Another perfectly equivalent way of defining effective subsidization is to say that the *social marginal internal rate of return to education,* $IRR_{soc} \equiv w_2 G' L_2/\left(f + w_1^s\right)$, falls short of the discount rate r. Whichever definition of effective subsidization one prefers, it differs from the conventional definition according to which education is subsidized when the cost of tuition is subsidized, $\varphi < f$. Focusing only on the cost of tuition when discussing the subsidization of higher education is, however, too restrictive and partial. The entire tax and transfer system affects the incentives to invest in education. Clearly, a negative value of Δ_E may result from subsidizing the cost of tuition. But even with no subsidy on tuition, education can be effectively subsidized because there are other components in the tax and transfer system that can result in negative values of Δ_E. For instance, reducing the tax on qualified labour increases the statutory return to education, w_2, and effectively subsidizes education. A tax on the return to savings, r, reduces the cost of education and works in the same direction. If the earnings function were concave and the taxpayer supplied non-qualified labour $L_1 – E > 0$ in the first period, another way of encouraging education would be to tax first-period non-qualified wage income, which reduces the cost of foregone earnings. However, if the earnings function is convex so that the taxpayer's optimization implies $L_1 = E$, there are no foregone earnings because the available time in period 1 is fully spent on education. Thus, there are only non-taxable costs of foregone leisure.

From our analysis and using equations (14) and (15), we finally get a *condition characterizing efficient education policy*:

$$\Delta_E = \rho\widehat{MC} = \rho\widehat{MR} = \rho\gamma_\gamma\widehat{E} \text{ with } \widehat{E} < 0. \tag{18}$$

An immediate implication from equation (18) is that it is efficient to subsidize education effectively, i.e., the effective wedge on education is negative, as the second-order elasticity of a Mincer-type earnings function, γ_γ, is positive.

PROPOSITION 2. *It is second best to subsidize education in effective terms.*

The pivotal role of the second-order elasticity for proposition 2 can be explained as follows. The definition, $\gamma_\gamma = \frac{E}{\gamma}\frac{d\gamma}{dE} = \frac{EG''}{G'} + 1 - \gamma$, reveals that efficient policy has to account for two interdependent effects. One is captured by the first term, EG''/G'. Obviously, convexity of the earnings function, $G'' > 0$, provides reason for subsidizing education. There is, however, a second and possibly countervailing effect, $1 - \gamma$, if the earnings function is elastic. By contrast, if the earnings function is inelastic, the case for a subsidy is strengthened. The less elastic the earnings function, the more education should be effectively subsidized.

Note that proposition 2 holds for any utility function. The utility function may be arbitrary except for the assumptions needed to guarantee that the planner's optimization is well behaved. This is an important insight, as results characterizing the efficient taxation of savings are less general. In the Ramsey model with finite periods, the question of whether it is efficient to tax savings or not critically depends on the choice of the utility function (Atkinson and Stiglitz 1972, Sandmo 1974). This is a remarkable difference, which can be explained as follows. Savings result in wealth generating capital income without requiring extra effort. By contrast, education enhances productivity. This increase in productivity results in higher income only if combined with labour, which requires additional effort. Hence, earning qualified labour income involves a double margin, educational choice and labour supply, while earning capital income does not. This difference explains and justifies differential taxation of saving and education.

The theoretical analysis produces an optimal policy rule: it is efficient to subsidize education in effective terms and there is an efficient subsidy level. While a comprehensive empirical analysis of the optimality rule is beyond the scope of the paper, it is inviting to look at the education policy of OECD countries and to check whether they pursue efficient education policies. This requires rewriting equation (18) and make is suitable for empirical analysis. The first step is to combine equations (18) and (17):

$$\rho\gamma_\gamma\widehat{E} = \Delta_E = \frac{w_2 G' L_2/r - \left(f + w_1^s\right)}{\varphi + \omega_1^s} = \frac{\left(\frac{w_2}{r} - \frac{\omega_2}{\rho}\right) G' L_2 - (f - \varphi) - \left(w_1^s - \omega_1^s\right)}{\omega_2 G' L_2/\rho}. \tag{19}$$

When translating the efficiency condition in an empirically testable condition, we have to acknowledge that not all variables in equation (19) are observable. In particular, the efficient reduction in education, $\widehat{E}$, is not observable, nor is the difference between the social and private costs of foregone leisure, $w_1^s - \omega_1^s$, whenever earnings functions are convex. The idea to separate observable from non-observable variables suggests the following notation. Let

$$PB \equiv \omega_2 G' E L_2/\rho$$

be the *private benefit* of education. Subtracting the direct cost yields the *net private benefit*:

$$NPB \equiv PB - \varphi E.$$

Similarly, we define *net government benefit*

$$NGB \equiv \left(\frac{w_2}{r} - \frac{\omega_2}{\rho}\right) G' E L_2 - (f - \varphi)E.$$

And the ratio at which net returns to education are shared between the government and the individual is the *net benefit sharing ratio*:

$$NBR \equiv \frac{NGB}{NPB}.$$

Optimal individual behaviour requires *NPB* to be equal to the indirect cost of education, which, in our simple model, is determined by the private cost of foregone non-qualified leisure, $\omega_1^s E$. However, translating the theory in an empirical framework requires a broader understanding of the indirect costs of education. Observed values of *NPB* are so high that it is implausible to assume that they reflect only the cost of leisure. And indeed, Heckman et al. (2006) and others point to the high risks of schooling decisions. There is the risk of failure because higher education cannot be successfully completed, higher education does not result in employment or individuals switch disciplines, which also involves costs. It is useful to think of those costs of risk and re-optimization as being indirect and convex, just as the cost of leisure is indirect and convex. Hence, we suggest interpreting the costs modelled in the theoretical section as an example of any indirect convex costs that determine the propensity to invest in education. Let

$$ICR \equiv \frac{w_1^s - \omega_1^s}{\omega_1^s}$$

be the *indirect cost ratio*, which is the ratio at which indirect costs of education are shared between the government and the individual. Using all these definitions, equation (19) can finally be rewritten as

$$NBR = ICR + \rho\gamma_\gamma \frac{PB}{NPB}\widehat{E}. \tag{20}$$

Our objective is to check whether and to what degree OECD countries pursue efficient education policies. For this purpose three notions of efficiency have to be kept apart. Unconstrained efficiency requires that $0 = \widehat{E} = NBR = ICR$. However, in a world with taxation, unconstrained efficiency cannot prevail and, therefore, is entirely irrelevant for policy analysis. By contrast, efficiency in the partial analytical sense might play an empirical role. It can be characterized by $NBR = ICR \neq 0$. Finally, second-best policy requires that $NBR < ICR$. This follows from equation (20) and the analysis of the preceding sections. As has been argued, the relative change in education, $\widehat{E}$, is negative in second best, and the second-order elasticity of the earnings function, γ_γ, is positive by assumption. Hence, the second term on the right-hand side of equation (20) is negative. The net benefit ratio, *NBR*, must be smaller than the indirect cost ratio, *ICR*, with a second-best policy of taxing labour and subsidizing education.

COROLLARY. *In second best, the indirect cost ratio, ICR, exceeds the net benefit sharing ratio, NBR.*

6. Second-best tertiary education policy: An empirical application

The empirical research on earnings determination has a positive-theoretical focus. It aims to estimate the effect of a policy intervention on the marginal internal rate of return to schooling. For a discussion of the challenges in estimating the causal effects of schooling see Carneiro et al. (2011). The present paper follows a different strategy. It is normative-theoretic in nature and derives the conditions of efficient education policy. In the following, we confront the theoretical findings with data and try to assess the relative efficiency of education policy.[3] Such an undertaking is, no doubt, ambitious. Hence, the

3 There is a literature dealing with the efficiency of public spending, an input in the education production function, on education output like tertiary educational attainment or PISA scores (a recent example is Canton et al. 2018). Our focus is different. We do not ask if countries are on the production frontier or waste resources, but we focus on the optimal policy mix.

following analysis can only serve to open the discussion. We do this by studying the education tax policies of OECD countries.

The data used are from various OECD publications and comprise the years between 2008 and 2013. Missing observations are linearly interpolated but not extrapolated. Table A1 in the appendix describes the variables and the data sources, while table A2 summarizes the data. The focus of our analysis is on *NBR*, the ratio at which net returns to education are shared between the government and the individual. The data for computing *NBR* are taken from the OECD data on net public and net private benefits from tertiary education for men. On average, *NBR* is 0.52 but there is quite some variation, from a low of 0.04 up to 1.15. As a proxy for the indirect cost ratio *ICR*, we use the marginal tax wedge associated with a suitably chosen marginal tax rate τ. This is suggested by the definition $ICR = (w_1^s - \omega_1^s)/\omega_1^s = \tau/(1-\tau)$, with τ being the shadow tax rate on foregone leisure. Among the potential candidates, we use the marginal tax wedge for an average income, single worker with no children and denote it by MTW_{100}. The marginal tax wedge is computed from the net personal marginal tax rate reported by the OECD. Employer contributions to social security and net transfers are accounted for. Because our model focuses on individual education decisions, i.e., young individuals, the marginal tax rates for single individuals with no children seem appropriate. Note that MTW_{100} is on average 0.61 and varies between a low of 0.27 and a high of 1.47. MTW_{100} is a reasonable proxy for *ICR* because workers who finished secondary schooling are neither low- nor high-income workers. However, as an alternative proxy for *ICR*, and to include workers with below-average incomes, we use $MTW_{67,100}$ (average of MTW_{67} and MTW_{100}). Here the average value is 0.56.

Clearly, the marginal tax wedge is an imperfect proxy for the indirect costs of tertiary education, which also include for instance the risk of an investment in education. Hence, we add further controls to better capture the indirect cost of tertiary education. For example, to account for the risk of unemployment and the supply of workers with completed tertiary education, we include the unemployment rate of individuals with tertiary education as well as the percentage of workers with tertiary education in the labour force. Again, there is substantial variation in the sample of OECD countries. The average unemployment rate for workers with tertiary education is 4.4% and varies between a low of 1.4% and a maximum of 14%. We also find high variation regarding the percentage of individuals with tertiary education, which is between 14% and 53% (average is 32%). The average ratio of private benefit to net private benefit is 1.04 but can be as high as 1.19. This points to differences in the private direct costs of getting a tertiary education. To control for the relative income position of the highly educated, the relative earnings of individuals with less than tertiary education is added. The average earnings premium for tertiary education is, at 53%, substantial and there are differences between the countries. To proxy the (economic) ability of private households to invest in education, we also include the private savings rate. Countries also differ in

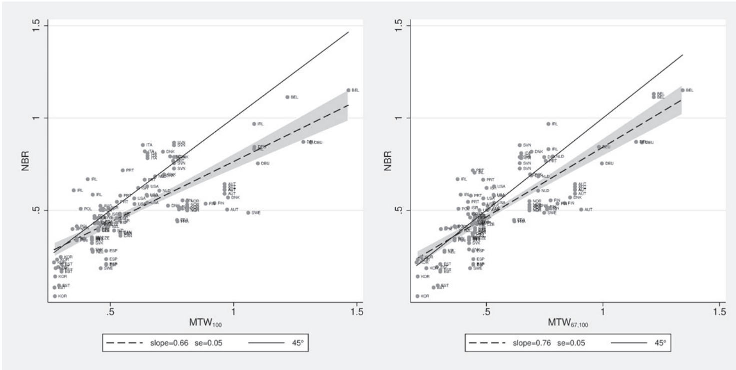

FIGURE 1 Correlation between NBR and MTW

the general quality of the education system; hence, we control for the countries' PISA 2009 and 2012 math scores. In fact, Hanushek and Wößmann (2010, 2015) have argued that performance on large-scale assessments is even a good predictor for future economic growth.

Besides variables to assess indirect costs of tertiary education, we need to account for the general inefficiency of public policy and the tax and transfer system that potentially affects the decision to invest in higher education. To control for political preferences for redistribution and taxation, which is typically associated with efficiency costs, we use the percentage of seats in parliament for social democratic parties as well as the Gini coefficient for disposable income. Moreover, the percentage of social expenditure from GDP describes how a country values and implements redistribution and is included as a proxy for inefficiency due to the redistribution of income. GDP growth rates and year dummies serve as a general measure of economic development. And finally, we exclude outliers from the following analyses. The criterion used is Cook's D.[4]

Second-best policy requires equation (20) to hold where, clearly, all three terms appearing in equation (20) are determined simultaneously. The left panel of figure 1 shows the scatter plot of *NBR* and the proxy MTW_{100} for *ICR*, and the right panel uses $MTW_{67,100}$. The first thing to note is the strong positive correlation between *NBR* and our alternative proxies MTW_{100} and $MTW_{67,100}$. Low values of *NBR* are found in Korea, whereas Belgium and Germany have high values indicating that the government strongly benefits from higher education. While it might be tempting to conclude from high values of *NBR* that more public support for tertiary education is needed, our

4 When using MTW_{100} as a proxy for *ICR*, we drop BEL in 2008, IRL in 2011 and ITA and PRT in 2008. If $MTW_{67,100}$ is used, the excluded observations are IRL in 2008 and 2013, ITA in 2008–2010, PRT in 2008 and SVN in 2011.

model points to the relationship between *NBR* and the tax wedges instead. The tax wedge and *NBR* are low in Korea and high in Germany and Belgium. Thus, policy conclusions based on *NBR* are in fact only misleading because they only partially account for variation in tax policy relevant for investment in higher education. A high level of *NBR* combined with a high level of MTW on the lower-educated might well be efficient.

From the theoretical analysis, we concluded that second-best policy requires countries to subsidize education. As argued in the preceding section, this means that the residual term in equation (20), $\rho\gamma_\gamma \frac{PB}{NPB}\widehat{E}$, is negative. Thus, all observations are expected to be on or below the 45° line. Observations above the line indicate inefficiency. And in fact, according to figure 1 the vast majority of observations are below the 45° line. However, being below the 45° line is only a first indicator for efficiency. Education policy might still be too generous or not generous enough. Figure 1 also shows the linear regression lines, with slopes being significantly less than one in both panels. The observations for Germany, for instance, are close on the regression line, indicating an average relationship between *NBR* and MTW. To get closer to the 45° line, Germany could either increase *NBR* or (and) decrease MTW. Acknowledging the mobility of high-skilled labour and the immobility of low-skilled labour in an open economy, the policy advice would be to decrease MTW, that is, to lower the net marginal tax rate on the less-educated. Note that this argument is not based on equity considerations but on results from an efficiency argument.

It is plausible to assume that the locus of efficiency is a curve passing through the origin and bending away from the 45° line for large values of MTW. This can be supported by the following reasoning: if *NBR* takes on large values, net returns to education are highly taxed, either because of high progressiveness in taxation or because the government budget requirements are generally high. In both cases, $-\widehat{E}$ should be high, which implies a large difference *NBR-ICR*. Moreover, because MTW is a proxy for *ICR* the distance to the 45° line should also be large. And clearly, if *NBR* takes on small values, $-\widehat{E}$ will be small. The regression line should come close to the 45° line in this case.

Figure 1 relies on the assumption that the marginal tax wedge on average non-skilled labour can be used to proxy the indirect cost ratio. This choice of proxy is suggested by the theoretical model that equates the indirect costs of education with the costs of foregone leisure. As noted, an empirical test of the theoretical analysis requires a broader interpretation of indirect costs. For instance, there is the risk of failure, which affects schooling decisions. Moreover, the inefficiency of the tax and transfer system needs to be accounted for. Hence, in a next step, we control for the additional indirect costs of tertiary education, the quality of the educational system and the inefficiency of the tax and transfer system by adding the control variables described in tables A1 and A2 as well as controlling for year fixed effects. On the basis of equation (20), we expect that the additional controls will result in a regression line that is closer to the 45° line than the earlier analyses without said controls. Now, the 45° line is to be interpreted as the locus of efficient educational

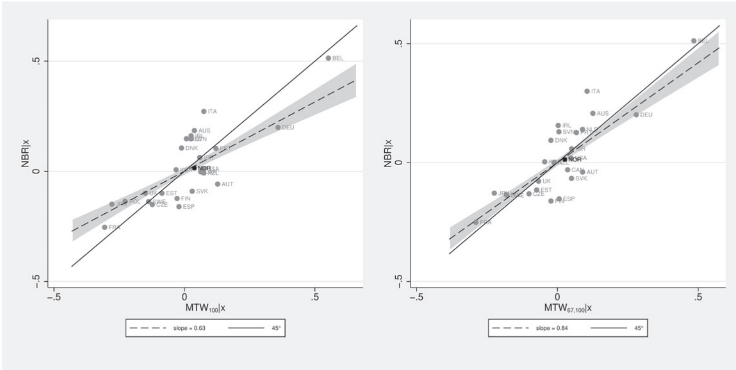

FIGURE 2 Correlation between NBR and MTW
NOTE: Controls included; scatter plots show country averages.

subsidy policies, provided all other distortions of the tax and transfer system have been controlled for. In figure 2, we do find a stronger relationship between $MTW_{67,100}$ and NBR, but the slope of the regression line, 0.84, is still significantly less than one (the figure shows country averages). Recall that in the regression analyses without controls, very few countries are above the 45° line. This changes after controlling for the general inefficiency of the tax and transfer system, which results in more countries being plotted in this group. For example, Italy and Australia are both above the 45° line and outside of the confidence interval, indicating that these countries not only subsidize education to a lesser degree than the average OECD country but also effectively tax tertiary education during the time period studied. Other countries like Spain and Austria subsidize tertiary education at a level that exceeds the subsidy levels of other countries in our sample.

One missing feature of the analysis is the lack of a policy benchmark. So far, we used the 45° line and the average policy position of OECD countries to assess individual country's educational policy. As a next step, we choose Norway as a policy benchmark. Norway is very close to the regression line and is a country for which $MTW_{67,100}$ is closest to NBR, after including the control variables. Figure 3 shows the estimated country effects and the confidence intervals obtained from a regression of the difference between NBR and MTW on our control variables, a set of country dummies, and without an intercept.

On average and over the time period studied, most countries do not exhibit a statistically significant difference in tertiary educational policy from the chosen benchmark. However, significant deviations from efficiency are estimated in some countries. Note, that a positive deviation from the benchmark suggests an inefficiently low subsidy or tax on higher education, whereas a negative deviation indicates inefficiently high subsidies. Figure 3 shows that in none of the countries in our sample is the level of higher educational subsidies too low. Rather, conditional on the set of control variables, subsidies in some

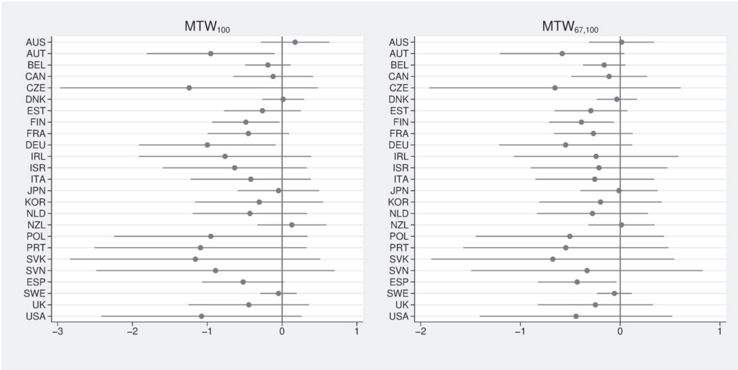

FIGURE 3 Deviations from efficiency
NOTE: Controls included; Norway is the benchmark; 95% confidence interval.

countries exceed the chosen benchmark. Some examples are Austria, Finland and Germany. Those countries have a high marginal tax wedge affecting the net income of workers without a higher education and working as an indirect subsidy for higher education. Spain, on the other hand, has comparably low MTW and NBR values. Thus, there is no common explanation for the observed deviations from the benchmark. This is not surprising given that our theoretical approach pointed to the challenge of finding the optimal subsidy for tertiary education in a second-best equilibrium and taking into account the various existing inefficiencies of the tax and transfer systems.

7. Conclusions

When labour supply is distorted by wage taxation, it is efficient to effectively subsidize education such that the marginal social cost of education exceeds the marginal social benefit. This result is noteworthy for two reasons. First, it is derived purely on grounds of efficiency; it does not draw on considerations of equity. While it is not surprising that equity considerations can justify subsidies to education, as has been convincingly argued by Bovenberg et al. (2005) and Krueger et al. (2013), the justification of subsidies to education based on efficiency arguments is not straight forward. Second, the efficiency-related justification for subsidizing education analyzed in the present paper does not rely on the assumption of market failure. The empirical evidence of externalities and liquidity constraints is mixed (Heckman et al. 1998, Lange et al. 2006, Carneiro et al. 2002). However, even if market failure were a valid concern, one could still only rationalize educational subsidies to the extent that the marginal social cost equals the marginal social benefit. Effective subsidization, as rationalized in the present paper, goes beyond this point.

The result is derived from a two-period model in which a utility maximizing representative household decides on labour and education in a setting wherein the government budget is financed by distortionary wage taxes. Subsidizing education is optimal because it alleviates the social cost of taxing qualified labour. In other words, a double margin requires effective subsidization of education. The key assumption driving this result is an earnings function that exhibits convexity and increasing elasticity. We call those functions Mincer-type earnings functions, and the Mincerian log-linear earnings function is just one example.

This paper also makes a methodological contribution to the literature. The Mincer–Becker assumption of an income maximizing schooling decision is replaced with utility maximization, as a broader and less restrictive concept. The two approaches are shown to have non-equivalent behavioural implications in the empirically relevant case of convex earnings functions. In the standard Mincer–Becker model of income maximization, optimality requires the marginal internal rate of return to schooling to equal the discount rate. The widely acknowledged problem with this characterization of optimality is that the marginal internal rate of return is notoriously estimated to exceed standard discount rates raising serious doubts about the rationality of the individuals' schooling decisions. We show that this inconsistency is resolved when assuming utility maximization instead of income maximization. Individuals do not fail to make rational choices; rather, empirical estimates of the internal rate of return systematically overstate their true values because they rely on foregone earnings as a proxy for the opportunity costs of education. Maximizing utility in a model with a convex earnings function, however, reveals that the true opportunity costs of education result from foregone leisure and that those costs systematically exceed the costs of foregone earnings. Hence, if the estimated internal rates of return to education are (seemingly too) high, the reasoning is systematic bias and not irrational choice.

We also confront theory with empirical evidence from OECD countries, which is complicated by the fact that key variables determining the choice of education are not directly observable. This is particularly true for the indirect costs of education. We solve this problem by using marginal tax wedges as proxies for the indirect cost ratio ICR, the ratio defining the government to individual indirect costs of higher education. Although the analysis is tentative, the results are interesting. It is shown that the vast majority of OECD countries effectively subsidize tertiary education. Compared with a benchmark, while no country subsidizes education at too low of a rate, there is evidence for sub-optimally high levels of subsidies to higher education. Sub-optimally high levels of subsidy occur when, e.g., the tax wedge for workers with less than tertiary education is too high. This finding adds a new perspective to the ongoing policy debate.

Appendix

See next page.

TABLE A1

Data description and sources

		Source
NBR	Net benefit sharing ratio: (total public benefit – direct public cost)/(total private benefit – direct private cost)	OECD (2012), tables A9.3 and A9.4 OECD (2013, 2014c, 2015, 2016, 2017), tables A7.3a and A7.4a
MTW_{100}	Marginal tax wedge for a single person at 100% of average earnings, no children	OECD, Social Expenditure Database (SOCX)
$MTW_{67,100}$	Average marginal tax wedge for a single person at 67 and 100% of average earnings, no children	OECD, Social Expenditure Database (SOCX)
Unemployed academics	Unemployment rates among men with tertiary education (%)	OECD (2014c, 2015), table A5.4a
Tertiary education	Share of the population with tertiary education, aged 25 to 64 (%)	OECD (2010a, 2011, 2013, 2014c), table A1.3a
Relative earnings	Relative earnings of 25- to 64-year-olds with income from employment (tertiary education)	OECD (2010a), table A7.1 OECD (2011, 2012), table A8.1 OECD (2012, 2013), table A6.1 OECD (2014c, 2015), table A6.1a
Private savings rate	Net private saving rate in percent of GDP	OECD (2014d), table 6.1
PISA math	PISA math score	OECD (2010b, 2014b)
Private ratio	Total private benefit / (total private benefit – direct private cost)	OECD (2012), table A9.1 OECD (2013, 2014c, 2015, 2016, 2017), table A7.1a
Social democrats	Share of seats for the party classified as a social democratic party (%)	Armingeon et al. (2019)
Social expenditure	Total public social expenditure in % of GDP	OECD, Social Expenditure Database (SOCX)
Gini coefficient	Gini-coefficient of disposable income	OECD, Income Distribution Database (IDD)
Real GDP growth rate	Real GDP growth rate	OECD (2014a), annex table 53

TABLE A2

Summary of the data, 2008–2013

	N	Mean	Std. dev.	Min.	Max.
NBR	133	0.52	0.22	0.04	1.15
MTW_{100}	133	0.61	0.25	0.27	1.47
$MTW_{67,100}$	133	0.56	0.22	0.20	1.34
Unemployed academics	133	4.40	2.33	1.37	14.01
Tertiary education	133	31.91	10.02	14.30	52.59
Relative earnings	133	152.99	18.73	121.20	185.67
Private savings rate	133	5.71	6.26	–8.27	27.75
PISA math	133	501.23	20.19	447	554
Private ratio	133	1.04	0.04	1.00	1.19
Social democrats	133	28.10	17.61	0.00	71.10
Social expenditure	133	21.93	5.19	7.65	31.49
Gini coefficient	133	0.30	0.04	0.24	0.40
Real GDP growth rate	133	2.28	3.67	–14.10	12.84

Supporting information

Supplementary material accompanies the online version of this article.

References

Anderberg, D. (2009) “Optimal policy and the risk-properties of human capital reconsidered,” *Journal of Public Economics* 93, 1017–26

Armingeon, K., V. Wenger, F. Wiedemeier, C. Isler, L. Knöpfel, D. Weisstanner, and S. Engler (2019) *Comparative Political Data Set 1960–2017*. Zurich: Institute of Political Science, University of Zurich. Available at `www.cpds-data.org/index.php/data`

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

Becker, G.S. (1964) *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. New York: National Bureau of Economic Research

Bovenberg, A.L., and B. Jacobs (2005) “Redistribution and learning subsidies are Siamese twins,” *Journal of Public Economics* 89, 2005–35

Canton, E., A. Thum-Thysen, and P. Voigt (2018) “Economists’ musings on human capital investment: How efficient is public spending on education in EU member states?,” European Economy Discussion Papers, no. 081

Card, D. (1999) “The causal effect of education on earnings.” In O. Ashenfelter and D. Card., eds., *Handbook of Labor Economics*, vol. 3, ch. 30, pp. 1801–63. Elsevier

Carneiro, P., and J. J. Heckman (2002) “The evidence on credit constraints in post-secondary schooling,” *Economic Journal* 112, 705–34

Carneiro, P., J. J. Heckman, and E. J. Vytlacil (2011) “Estimating marginal returns to education,” *American Economic Review* 101, 2754–81

Findeisen, S., and D. Sachs (2015) “Designing efficient college and tax policies,” CESifo working paper no. 5435

Hanushek, E.A., and L. Wößmann (2010) "Education and economic growth." In P. Peterson, E. Baker, and B. McGaw, eds., vol. 2, pp. 245–52. *International Encyclopedia of Education*. Oxford, UK: Elsevier
—— (2015) *The Knowledge Capital of Nations: Education and the Economics of Growth*. Cambridge, MA: The MIT Press
Heckman, J.J., and P. J. Klenow (1998) "Human capital policy." In M. Boskin, ed., *Policies to Promote Human Capital Formation*. Hoover Institution
Heckman, J.J., L. J. Lochner, and P. E. Todd (2006) "Earnings functions, rates of return and treatment effects: The Mincer equation and beyond." In E. A. Hanushek and F. Welch, eds., *Handbook of the Economics of Education*, vol. 1, ch. 7, pp. 307–458. Elsevier
—— (2008) "Earnings functions and rates of return," *Journal of Human Capital* 2, 1–31
Jacobs, B., and A. L. Bovenberg (2011) "Optimal taxation of human capital and the earnings function," *Journal of Public Economic Theory* 13, 957–71
Krueger, D., and A. Ludwig (2013) "Optimal progressive labor income taxation and education subsidies when education decisions and intergenerational transfers are endogenous," *American Economic Review* 103, 496–501
Lange, F., and R. Topel (2006) "The social value of education and human capital." In E. A. Hanushek and F. Welch, eds., *Handbook of the Economics of Education*, vol. 1, ch. 8, pp. 459–509. Elsevier
Lemieux, T. (2006) "The 'Mincer Equation' thirty years after *Schooling, Experience, and Earnings*." In S. Grossbard, ed., *Jacob Mincer: A Pioneer of Modern Labor Economics*, pp. 127–45. Boston, MA: Springer
Mincer, J. (1958) "Investment in human capital and personal income distribution," *Journal of Political Economy* 66, 281–302
Mirrlees, J.A. (1971) "An exploration in the theory of optimum income taxation," *Review of Economic Studies* 38, 175–208
OECD (2010a) *Education at a Glance 2010: OECD Indicators*. Paris: OECD Publishing
—— (2010b) *PISA 2009 Results: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science, Volume I*. Paris: OECD Publishing
—— (2011) *Education at a Glance 2011: OECD Indicators*. Paris: OECD Publishing
—— (2012) *Education at a Glance 2012: OECD Indicators*. Paris: OECD Publishing
—— (2013) *Education at a Glance 2013: OECD Indicators*. Paris: OECD Publishing
—— (2014a) *OECD Economic Outlook, Volume 2014, Issue 1*. Paris: OECD Publishing
—— (2014b) *PISA 2012 Results: What Students Know and Can Do (Volume I, Revised Edition, February 2014): Student Performance in Mathematics, Reading and Science, PISA*. Paris: OECD Publishing
—— (2014c) *Education at a Glance 2014: OECD Indicators*. Paris: OECD Publishing
—— (2014d) *National Accounts at a Glance 2014*. Paris: OECD Publishing
—— (2015) *Education at a Glance 2015: OECD Indicators*. Paris: OECD Publishing

—— (2016) *Education at a Glance 2016: OECD Indicators*. Paris: OECD Publishing

—— (2017) *Education at a Glance 2017: OECD Indicators*. Paris: OECD Publishing

OECD Income Distribution Database (IDD). Available at www.oecd.org/social/income-distribution-database.htm

OECD Social Expenditure Database (SOCX). Available at www.oecd.org/social/expenditure.htm

Oreopoulos, P., and K. G. Salvanes (2011) "Priceless: The nonpecuniary benefits of schooling," *Journal of Economic Perspectives* 25(1), 159–84

Ramsey, F.P. (1927) "A contribution to the theory of taxation," *Economic Journal* 37, 47–61

Richter, W.F. (2009) "Taxing education in Ramsey's tradition," *Journal of Public Economics* 93, 1254–60

Roy, A. (1951) "Some thoughts on the distribution of earnings," *Oxford Economic Papers* 3, 135–46

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

Wallenius, J. (2011) "Human capital accumulation and the intertemporal elasticity of substitution of labor: How large is the bias?," *Review of Economic Dynamics* 14, 577–91

Willis, R.J., and S. Rosen (1979) "Education and self-selection," *Journal of Political Economy* 87, S7–S36

Zimmerman, S.D. (2014) "The returns to college admission for academically marginal students," *Journal of Labor Economics* 32, 711–54