TU Dortmund University

Education: Optimal Choice and Efficient Policy

by

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

and

Kerstin Schneider
WIB University of Wuppertal and CESifo Munich

This version: November 2019

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 distorting 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.

JEL Classification: (H21, I28, J24)

Keywords: schooling choice; utility vs. earnings maximization; convex earnings functions; second-best taxation in Ramsey’s tradition; empirical evidence for OECD

Address:
Wolfram F. Richter, TU Dortmund University, Department of Economics, 44221 Dortmund, Germany. E-mail: Wolfram.Richter@tu-dortmund.de; Phone: +49 231 755 3146. Kerstin Schneider, University of Wuppertal, Wuppertal Research Institute for the Economics of Education, 42119 Wuppertal, Germany. E-mail: Schneider@wiwi.uni-wuppertal.de; Phone: +49 202 439 2483.

1 This is a fully revised and extended version of a paper written by the first author and previously circulating under the title “Mincer equation, power law of learning, and efficient education policy”. Helpful comments of the editor, Arthur Sweetman, and anonymous referees are gratefully acknowledged.
1. Introduction

Growth and welfare in the knowledge society depends on countries’ investments in human capital; and since 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-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 modeling 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 behavior is challenged by evidence of significant nonpecuniary 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 nonpecuniary returns, Oreopoulos et al. (2011, p. 180) conclude that nonpecuniary returns “are both real and important”. As to the nonpecuniary costs of education, Heckman et al. (2006, p. 436) suggest that psychic costs “play a
very important role”. However, the black-box character of nonpecuniary 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 behavior and it also serves as a basis for the analysis of allocational efficiency. Initially, utility and income maximization appear to be concepts with equivalent behavioral 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 which 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, one, however, has to deal with corner solutions which 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 which is log-linear in schooling like the standard Mincer earnings function features 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 or distributional issues, 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 labor 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), Krueger et al. (2013). A major shortcoming of this literature is that it assumes concave earnings functions despite empirical evidence from the literature on labor economics reinforcing the convexity of the earnings function. In fact, there is

---

2 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.
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 modeled 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 only maximize income when making schooling decisions. They suggest accounting for heterogeneity and including psychic costs in the analysis. As compared to 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 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, since 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 labor 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 modeling 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 labor 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,

$$\ln G(E) = a + mE,$$

(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 either invest in education or supply labor but they never do both in the same period of life. The consequence of increasing elasticity is that it is second-best efficient to subsidize education labor supply is distorted by wage taxes.

4. Household behavior

In the following we introduce the Mincer-type earnings function in a standard model of household behavior 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, $L_i \leq \bar{L}_i$, in periods $i=1,2$. $\bar{L}_i$ 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 labor supply; whereas in period 1, $L_1 - E$ is labor supply with $E$ being time spent on education. First-period labor earns a constant wage rate after tax, $\omega_1$; the return to second-period labor, 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 $\omega_2$ accounts for wage taxation. I.e. $\omega_2$ equals one minus the wage tax of period 2. Given a positive choice of education, $E > 0$, second-period labor is interpreted as qualified labor. Likewise, the quantities $L_1 - E$ and $L_1$ are interpreted as nonqualified labor and nonqualified 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 modeled by $\omega_1 E$, and the cost of tuition is $\varphi E$. The share of first-period income that is not spent on education or on consumption is first-period savings:

$$S = \omega_1 (L_1 - E) - \varphi E - C_1 = \omega_1 L_1 - (\varphi + \omega_1)E - C_1$$ (2)
By way of normalization, the price of consumption is set equal to one. The gross rate of return to saving is denoted by \( \rho \), 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:

\[
C_2 = \rho S + \omega_2 G(E)L_2
\]  
(3)

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

\[
C_1 + \frac{C_2}{\rho} = \omega_1 L_1 + \omega_2 G(E)L_2/\rho - (\varphi + \omega_1)E
\]  
(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 (4), \( \bar{L}_1 \geq L_1 \geq E \), and \( \bar{L}_2 \geq L_2 \) requires that net income, \( Y \), is maximized in \( E \) holding other variables constant,

\[
Y(L_1, L_2) = Y(L_1, L_2; \omega_2/\rho(\varphi + \omega_1))
\]

\[
\equiv \max_{0 \leq E \leq L_1} [\omega_2 G(E)L_2/\rho - (\varphi + \omega_1)E] .
\]  
(5)

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

When maximizing (5), three scenarios are conceivable. 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 labor 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 only happen 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 either on unqualified labor supply or on 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,

$$IRR_{priv} \equiv \frac{\omega_2 G'(E)L_2}{\varphi + MRS^1} = \rho.$$  \hspace{1cm} (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*, $\omega_1$. It has, however, been shown that with utility maximizing individuals this equality only holds 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 $L_2$:

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

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

$$\frac{d^2 Y}{d L_1^2} = \frac{\omega_2}{\rho} G'' L_2 > 0. \quad (7.b)$$

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 $L_i \in [0, \bar{L}_i]$ may fail to exist. Still, the following analysis only looks at first-order conditions. The implicit assumption is, firstly, that the individual discards all solutions of the first-order conditions which fail to be globally optimal and, secondly, that Inada-type conditions hold. The latter implies that marginal disutility of non-leisure tends to infinity when $L_i$ approaches the upper bound, $\bar{L}_i$, 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, $G = G(E_1 + \ldots + E_i)$. 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 $\bar{t}$ after which individuals switch from education to work. Thus individuals find it optimal to be educated in all earlier periods, $i < \bar{t}$, and to work in all later periods, $i \geq \bar{t}$. Allowing for heterogeneous individuals, the cut-off period $\bar{t}$ 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 Appendix A of the Online Appendices.)

5. Second-best policy

We now turn to optimal policy design in the two-period model, keeping in mind that the results equally apply 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 labor income in the first and the second period, on the cost of tuition, and on the returns to savings. They are modeled 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 $w_1, w_2, f, r$. The tax on labor 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

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

3 The function $G(E)$ has been introduced as earnings, suggesting $w_2 = 1$. If one chose instead to interpret education as a labor augmenting activity and $G(E)L_2$ as effective qualified labor, $w_2$ equaled 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.
In order to characterize second-best tax policy, it is convenient to work with the taxpayer’s expenditure function which is defined by

\[ e(\omega_1, \omega_2, \varphi, \rho; u) \equiv \min[\rho C_1 + C_2 + \rho(\varphi + \omega_1)E - \rho\omega_1L_1 - \omega_2G(E)L_2] \tag{9} \]

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 \( \omega_1 \), 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, eq. (8) can be written as

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

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,

\[ \hat{E} = \hat{L}_1 = \hat{G}L_2 = \hat{C}_1. \tag{11} \]

A hat on a function \( X = X(\omega_1, \omega_2, \varphi, \rho; u) \) denotes a relative change, \( \hat{X} \equiv \Delta X/X \), with the total differential operator \( \Delta \) 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 (12), $\Delta X$ 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

$$\widehat{C}_2 = \hat{E}, \widehat{L}_2 = (1 - \gamma)\widehat{L}_1, \text{ and } \widehat{G} = \gamma\hat{E},$$

(13)

where $\gamma$ 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 Appendix B of the Online Appendices.)

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 labor supply, $L_2$. Efficient taxation requires the relative reduction in qualified labor to be smaller than the relative reduction in non-qualified labor. The factor is $1 - \gamma$ and hence, the tax induced reduction in qualified labor decreases in $\gamma$. Thus, the more elastic the individual earnings function is, the less the optimal reduction in qualified labor relative to nonqualified labor. 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 labor, $G\widehat{L}_2$, is reduced equi-proportionately. Qualified labor, however, should be reduced less than proportionately, as the elasticity of the earnings function $\gamma$ is positive. For earnings functions with elasticity $\gamma$ greater than one, $L_2$ should even increase (cf. eq. (13)).

The optimal choice of education is characterized in eq. (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,
\[
MR \equiv G'(E)L_2 = \rho(\varphi + MRS^1)/\omega_2 \equiv MC. \tag{14}
\]

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

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

where \(\gamma := E \gamma' / \gamma\) 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 \(E\). \(\gamma\) equals one if the earnings function is log-linear (the Mincerian case). As \(\hat{E}\) is negative, given the need to raise positive tax revenue, it follows from (15) that the efficient change in the marginal return to education, \(\hat{MR}\), is negative as well. Since \(MR\) 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 eq. (14) yields

\[
0 > \hat{MC} = \frac{\Delta MC}{MC} = \frac{\Delta(\rho \varphi) + \Delta(\rho MRS^1)}{\rho(\varphi + MRS^1)} - \frac{\Delta \omega_2}{\omega_2} - \frac{\omega_2 - w_2}{r \omega_2}
= \frac{w_2 / r}{\omega_2} - \frac{f + MRS^1 - \rho \Delta MRS^1}{\rho(\varphi + MRS^1)} \cdot \tag{16}
\]

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 \(\omega_1\) only if the earnings function is concave. If the earnings function is convex, it is nevertheless suggestive to write \(MRS^1 \equiv w_1^s\) and \(MRS^1 - \rho \Delta MRS^1 \equiv w_1^s\) and to interpret \(w_1^s\) and \(w_1^s\) as the respective private and social shadow costs of foregone nonqualified leisure. Hence, eq. (16) can be restated as

\[
0 > \rho \hat{MC} = \frac{w_2 / r}{\omega_2 / \rho} - \frac{f + w_2^s}{\rho w_1^s} = \frac{w_2 G' L_2 / r - (f + w_2^s)}{\varphi + w_1^s} \equiv \Delta_E, \tag{17}
\]
where $\Delta_E$ is interpreted as the \textit{effective wedge on education}, and eq. (6) has been used. The inequality in (17) is equivalent to $w_2G'L_2/r < f + w_1^s$.

In the following we refer to \textit{education} as (\textit{effectively}) \textit{subsidized} when the effective wedge on education is negative. A negative wedge requires a level of education for which the (\textit{effective}) \textit{marginal social cost}, $r(f + w_1^s)/w_2$, exceeds the \textit{marginal social return}, $G'L_2$. Another perfectly equivalent way of defining effective subsidization is to say that the \textit{social marginal internal rate of return to education}, $IRR_{soc} \equiv w_2G'L_2/(f + w_1^s)$, 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, $\rho < f$. 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 $\Delta_E$ 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 $\Delta_E$. For instance, reducing the tax on qualified labor 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 nonqualified labor $L_1 - E > 0$ in the first period, another way of encouraging education would be to tax first-period nonqualified 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 as 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 eq. (14) and eq. (15) we finally get a condition characterizing efficient education policy:

$$\Delta_E = \rho \bar{MC} = \rho \bar{MR} = \rho \gamma_E \bar{E} \quad \text{with} \quad \bar{E} < 0 .$$  (18)

An immediate implication from eq. (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, $\gamma$, 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, $$\gamma = \frac{E}{\gamma} \frac{d\gamma}{dE} = \frac{EG''}{G'} + 1 - \gamma ,$$ 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 is, the more should education 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 only in higher income if combined with labor, which requires additional effort. Hence earning qualified labor income involves a double margin, educational choice and labor 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 eq. (18) and make it suitable for empirical analysis. The first step is to combine equations (18) and (17):

\[ \rho \gamma_f \dot{E} = \Delta_E = \frac{w_2 G' L_2 / \tau - (f + w_1^s)}{\varphi + w_1^s} = \frac{(w_2 - \omega_2) G' L_2 - (f - \varphi) - (w_1^s - \omega_1^s)}{\omega_2 G' L_2 / \rho} \].

(19)

When translating the efficiency condition in an empirically testable condition, it has to be acknowledged that not all variables in eq. (19) are observable. In particular, the efficient reduction in education, \( \dot{E} \), is not observable, neither 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*
\[ NGB \equiv \left( \frac{\omega_2}{r} - \frac{\omega_2}{\rho} \right) G'EL_2 - (f - \varphi)E. \]

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 behavior requires \( NPB \) to be equal to the indirect cost of education which, in our simple model, is determined by the private cost of foregone nonqualified leisure, \( \omega_1^sE \). 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 only reflect 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 either because higher education cannot be successfully completed, 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 modeled in the theoretical section as an example of any indirect convex costs which determine the propensity to invest in education. Let

\[ ICR \equiv \frac{w_i^s-\omega_1^s}{\omega_1^s} \]

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, eq. (19) can finally be rewritten as

\[ NBR = ICR + \rho \gamma \frac{PB}{NPB} \hat{E}. \] (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 \( 0 = \hat{E} = NBR = ICR \). However, in a world with taxation, unconstrained efficiency cannot prevail and hence 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 eq. (20) and the analysis of the preceding sections. As has been argued, the relative change in education, \( \hat{E} \), is negative in second-best and the second-order elasticity of the earnings function, \( \gamma_\gamma \), is positive by assumption. Hence the second term on the right-hand side of eq. (20) is negative. The net benefit ratio, \( NBR \), must be smaller than the indirect cost ratio, \( ICR \), with a second-best policy of taxing labor and subsidizing education.

Corollary: In second-best, the indirect cost ratio, \( ICR \), exceeds the net benefit sharing ration, \( NBR \).


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.\(^4\) Such an undertaking is, no doubt, ambitious. Hence, the following analysis can only serve to open the discussion. We do this by studying the education tax policies of OECD countries.

The data used is from various OECD publications and comprises the years between 2008 and 2013. Missing observations are linearly interpolated but not extrapolated. Table A.1 in the Appendix describes the variables and the data sources, while Table A.2 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 \(\tau\). This is suggested by the definition \(ICR = (w_1^s - \omega_1^s)/\omega_1^s = \tau/(1 - \tau)\), with \(\tau\) 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. Since 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

---

\(^4\) 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, Thum-Thysen, Voigt, (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.
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 labor force. Again, there is substantial variation in the sample of OECD countries. The average unemployment rate for workers with tertiary education is 4.4 percent and varies between a low of 1.4 and a maximum of 14 percent. We also find high variation regarding the percentage of individuals with tertiary education, which is between 14 and 53 percent (average is 32 percent). 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 percent, 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 the general quality of the education system, hence we control for the countries’ PISA 2006, 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.$^5$

Second-best policy requires eq. (20) to hold where, clearly, all three terms appearing in eq. (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 model points to the relationship between $NBR$ and the tax wedges instead. The tax wedge as well as $NBR$ are low in Korea and high in Germany and Belgium. Thus, policy conclusions based on $NBR$ only are in fact misleading, as 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.

$^5$ 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.
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 eq. (20), \( \rho \gamma \frac{PB}{NPB} \hat{E} \), 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 \( NBR \) and/or decrease \( MTW \). Acknowledging the mobility of high skilled and the immobility of low skilled labor 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 results from an efficiency argument.

– Figure 1 about here –

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 \( -\hat{E} \) 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, \( -\hat{E} \) 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 labor 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 A.1 and A.2 as well as controlling for year fixed effects. Based on eq. (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 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, 0.82, 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 they also effectively tax tertiary education during that 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.

– Figure 2 about here –
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 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.

– Figure 3 about here –

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 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, as 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 labor 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 straightforward. 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 labor 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 labor. 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 behavioral 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. And while, compared to a benchmark, 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

Table A1. Data description and sources

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net benefit sharing ratio</td>
<td><em>(men)</em></td>
<td>OECD: Education at a Glance</td>
</tr>
<tr>
<td>$MTW_{100}$</td>
<td>Marginal tax wedge for a single person at 100% of average earnings, no children</td>
<td>OECD iLibrary</td>
</tr>
<tr>
<td>$MTW_{67,100}$</td>
<td>Average marginal tax wedge for a single person at 67 and 100% of average earnings, no children</td>
<td>OECD iLibrary</td>
</tr>
<tr>
<td>Unemployed academics</td>
<td>Unemployment rates among men with tertiary education (%)</td>
<td>OECD: Education at a Glance</td>
</tr>
<tr>
<td>Tertiary education</td>
<td>Share of the population with tertiary education, aged 25-64 years (%)</td>
<td>OECD: Education at a Glance</td>
</tr>
<tr>
<td>Relative earnings</td>
<td>Relative earnings of 25-64 year-olds with income from employment (tertiary education)</td>
<td>OECD: Education at a Glance</td>
</tr>
<tr>
<td>Private savings rate</td>
<td>Net private saving rate in percent of GDP</td>
<td>National Accounts at a Glance</td>
</tr>
<tr>
<td>Private benefit</td>
<td>Total private benefit</td>
<td>OECD: Education at a Glance</td>
</tr>
<tr>
<td>Net private benefit</td>
<td>Total private benefit - direct private cost</td>
<td>OECD: Education at a Glance</td>
</tr>
<tr>
<td>Public benefit</td>
<td>Total public benefit</td>
<td>OECD: Education at a Glance</td>
</tr>
<tr>
<td>Net public benefit</td>
<td>Total public benefit - direct public cost</td>
<td>OECD: Education at a Glance</td>
</tr>
<tr>
<td>Social democrats</td>
<td>Share of seats for the party classified as a social democratic party (%)</td>
<td>The Comparative Political Data Set III 1990-2015 (University of Bern)</td>
</tr>
<tr>
<td>Social expenditure</td>
<td>Total public social expenditure in % of GDP</td>
<td>Social issues: Key tables from OECD</td>
</tr>
<tr>
<td>Gini coefficient</td>
<td>Gini-coefficient of disposable income</td>
<td>OECD iLibrary</td>
</tr>
<tr>
<td>Real GDP growth rate</td>
<td></td>
<td>OECD Economic Outlook</td>
</tr>
</tbody>
</table>
Table A2. Summary of the data, 2008-2013

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBR(men)</td>
<td>133</td>
<td>0.52</td>
<td>0.22</td>
<td>0.04</td>
<td>1.15</td>
</tr>
<tr>
<td>$MTW_{100}$</td>
<td>133</td>
<td>0.61</td>
<td>0.25</td>
<td>0.27</td>
<td>1.47</td>
</tr>
<tr>
<td>$MTW_{67,100}$</td>
<td>133</td>
<td>0.56</td>
<td>0.22</td>
<td>0.20</td>
<td>1.34</td>
</tr>
<tr>
<td>Unemployed academics</td>
<td>133</td>
<td>4.40</td>
<td>2.33</td>
<td>1.37</td>
<td>14.01</td>
</tr>
<tr>
<td>Tertiary education</td>
<td>133</td>
<td>31.91</td>
<td>10.02</td>
<td>14.30</td>
<td>52.59</td>
</tr>
<tr>
<td>Relative earnings</td>
<td>133</td>
<td>152.99</td>
<td>18.73</td>
<td>121.20</td>
<td>185.67</td>
</tr>
<tr>
<td>Private savings rate</td>
<td>133</td>
<td>5.71</td>
<td>6.26</td>
<td>-8.27</td>
<td>27.75</td>
</tr>
<tr>
<td>PISA Math</td>
<td>133</td>
<td>501.85</td>
<td>19.58</td>
<td>446.86</td>
<td>546.23</td>
</tr>
<tr>
<td>Private ratio</td>
<td>133</td>
<td>1.04</td>
<td>0.04</td>
<td>1.00</td>
<td>1.19</td>
</tr>
<tr>
<td>Social democrats</td>
<td>133</td>
<td>28.10</td>
<td>17.61</td>
<td>0.00</td>
<td>71.10</td>
</tr>
<tr>
<td>Social expenditure</td>
<td>133</td>
<td>21.93</td>
<td>5.19</td>
<td>7.65</td>
<td>31.49</td>
</tr>
<tr>
<td>Gini coefficient</td>
<td>133</td>
<td>0.30</td>
<td>0.04</td>
<td>0.24</td>
<td>0.40</td>
</tr>
<tr>
<td>Real GDP growth rate</td>
<td>133</td>
<td>2.28</td>
<td>3.67</td>
<td>-14.10</td>
<td>12.84</td>
</tr>
</tbody>
</table>

References


Figure 1: Correlation between NBR and MTW
Controls included; scatter plots show country averages

*Figure 2:* Correlation between NBR and MTW, controls included
Figure 3: Deviations from Efficiency