A CASE FOR TAXING EDUCATION

TOMER BLUMKIN
EFRAIM SADKA

CESIFO WORKING PAPER NO. 1440
CATEGORY 1: PUBLIC FINANCE
APRIL 2005

CESifo *Working Paper No. 1440*

A CASE FOR TAXING EDUCATION

Abstract

We illustrate a novel informational feature of education, which the government may utilize. Discretionary decisions of individuals to acquire education may serve as an additional signal (to earned labor income) on the underlying unobserved innate earning ability, thereby mitigating the informational constraint faced by the government. We establish a case for taxing education, as a supplement to the labor income tax.

JEL Code: H2, D6.

Keywords: optimal taxation, re-distribution, education, inequality.

Tomer Blumkin
Department of Economics
Ben-Gurion University
Beer-Sheba 84105
Israel
tomerblu@bgumail.bgu.ac.il

Efraim Sadka
The Eitan Berglas School of Economics
Tel Aviv University
Tel-Aviv 69978
Israel
sadka@post.tau.ac.il

1. Introduction

Government provision and/or subsidization of education is often warranted on efficiency grounds, due to the existence of market failures, such as capital market imperfections [see, e.g., Barham et al. (1995)], moral hazard issues [see, e.g., Wilson (1999)], time inconsistency problems [see, e.g., Boadway et al. (1996) and Gradstein (2000)] and externalities [see, e.g., Eckstein and Zilcha (1994)]. The increase in earnings inequality over the last two decades is often attributed to the rise in the returns to education, thus rendering ever so relevant the debate on the role of publicly provided education in pursuing distributional goals.[1] When assessing the redistributive role of publicly provided education (as is often the case at the elementary, secondary, and post-secondary levels) or subsidized privately acquired education (most relevant at post secondary level), one has to take into account the existence of other redistributive fiscal instruments; notably, the labor income tax system. Indeed the optimal tax literature has examined the productivity enhancing role of education, alongside redistributive taxation in settings with informational asymmetries [see, e.g., Sheshinski (1971), Ulph (1977) and Tuomala (1986) for early contributions; and more recently, Boadway and Marchand (1995) and Bret and Weymark (2003)]. These studies differ in the way the government role in the market has been modeled. In some studies, education is not a policy tool [see, e.g., Sheshinski (1971)]; while in others [see, e.g., Boadway and Marchand (1995)], compulsory publicly provided education is examined; and in still others [see, e.g., Bret and Weymark (2003)], subsidizing discretionary investment in education is analyzed.

PT[1] **TP**Correcting market failures need not necessarily stand in conflict with redistributive purposes. Thus, for instance, alleviating credit constraints may be pro-poor if, plausibly, the incidence of credit problems is higher amongst individuals coming from disadvantaged socio-economic backgrounds.

In this paper we attempt to illustrate a novel informational feature of education, which the government may utilize: discretionary decisions of individuals to acquire education may serve as a supplementary signal (to earned labor income) on the underlying unobserved innate earning ability, thereby mitigating the informational constraint faced by the government.[2] Notably, contrary to conventional wisdom, employing a generalized version of the original model of Mirrlees (1971), where individuals decide whether to acquire productivity-augmenting education, we illustrate that a case for taxing education, as a supplement to the labor income tax, can be established.

The structure of the remainder of the paper is as follows. In the coming section we present the model. In section 3 we analyze the optimal policy. We conclude in section 4.

2. The Model

Consider an economy with a continuum of individuals (whose number is normalized to one), producing a single consumption-good. The production technology employs labor only and exhibits constant returns to scale and perfect substitutability. Following Bret and Weymark (2003), we assume that individuals differ in two characteristics: the innate ability, denoted by *w*, and scholastic aptitude, which is given by the cost of acquiring education (in forgone consumption terms), denoted by *e*. For simplicity, we assume that for a proportion $0 < \gamma_1 < 1$ of the individuals (referred to as type 1), the cost of education is given by e_1, whereas the innate ability

PT[2] **TP** The notion of using an education tax in a second best environment has been alluded to by Bret and Weymark (2003). The mechanism at work while employing such an education tax is closely related to the rationale for using commodity taxation as a supplement to income taxation, when additional information on unobserved innate ability can be inferred from variation in consumption patterns across individuals [see Atkinson and Stiglitz (1976), Deaton (1979) and (1981), inter-alia, on this matter]. In this sense, education expenditure is analogous to consumption choice.

is distributed according to a cumulative distribution function $G_1(w)$, with w denoting innate ability. The cost of education and the cumulative distribution function of the innate ability of type-2 individuals (who constitute a fraction $\gamma_2 = 1 - \gamma_1$ of the population) are given, respectively, by e_2 and $G_2(w)$. Both G_1 and G_2 have strictly positive densities and the same support - $[\underline{w}, \overline{w}]$. For concreteness we assume that $e_2 > e_1$. The cost-of-education and the innate ability characteristics are assumed to be private information.

We follow Saint-Paul (1994) and Razin and Sadka (2001), by assuming that the productivity of an individual of ability *w*, who acquires education, is given by *aw*, where *a*>1. We denote by z the productivity of an individual, where $z = aw$, if an individual acquires education and $z = w$, otherwise. The productivity of an individual is also the wage rate she earns.

All individuals share the same preferences given by a quasi-linear utility function:

(1) $U(c,l) = c - h(l)$,

where *c* denotes consumption, *l* denotes labor and *h*(*l*) is increasing, strictly convex and twice continuously differentiable.[3]

We assume that a linear tax system is in place, where the marginal tax rate is denoted by *t*, and the uniform lump-sum transfer (possibly negative) is given by τ. A typical individual with characteristics (w, e) has two kinds of decision to make. She has to make a binary choice whether to acquire education or not (this will be formally given by an indicator function *K*, where *K*=1, if she acquires education and *K*=0 otherwise); and she has to determine her labor supply. Consider first the labor supply

PT[3] **TP**Including the right-derivative at $l = 0$.

choice. An individual with productivity z is seeking to maximize the utility function in (1) subject to the following budget constraint:

(2) $$(1-t)zl+\tau=c\,.$$

Substituting the maximizing levels of consumption $[c(z,t,\tau)]$ and labor $[l(z,t,\tau)]$ into the utility function in (1), one obtains an indirect utility function denoted by $V(z,t,\tau)$. Turing next to the education choice, recalling the quasi-linear specification of the utility function, and plausibly assuming that education costs are non tax deductible,[4] an individual with characteristics (w,e) acquires education if and only if the following condition holds:

(3) $$V(aw,t,\tau)-V(w,t,\tau)\geq e\,.$$

It is straightforward to verify that:

(4) $$\frac{\partial[V(aw,t,\tau)-V(w,t,\tau)]}{\partial w}=(1-t)\cdot[a\cdot l(aw,t,\tau)-l(w,t,\tau)]>0\,,$$

because l is strictly increasing with respect to productivity. We conclude that for each type j (j=1,2), there exits a cutoff level of innate ability, $\hat{w}_j$, which is given by the implicit solution to:

(5) $$V(a\hat{w}_j,t,\tau)-V(\hat{w}_j,t,\tau)=e_j\,,\;\; j=1,2\,,$$

such that individuals acquire education if and only if their innate ability lies above this cutoff level. We plausibly assume that the cutoff (for each type) is bounded away from the lower bound of the support of innate abilities ($\hat{w}_j > \underline{w}$, $j=1,2$).

PT[4] **TP** Note that the cost of education may be associated with both the pecuniary costs of education and the opportunity cost of forgoing labor time. The latter may be non deductible if we, plausibly, assume that students' incomes lie below the tax-threshold; namely, the minimum level of earnings above which individuals pay taxes.

3. The Optimal Policy

The government is seeking to maximize some egalitarian social welfare function by choosing the tax instruments, t and τ, subject to a revenue constraint, taking into account the optimal choices of the individuals. The social planner is assumed to have the following objective function:

(6) $$W \equiv \sum_{j=1}^{2} \gamma_j \left[\int_{\hat{w}_j}^{\bar{w}} W[V(aw,t,\tau) - e_j] dG_j(w) + \int_{\underline{w}}^{\hat{w}_j} W[V(w,t,\tau)] dG_j(w) \right]$$

where $\hat{w}_j$ is given by equation (5), j=1,2, and where $W'(V) > 0$ and $W''(V) < 0$, thus the welfare function is exhibiting strict inequality aversion. The objective in (6) is maximized subject to the government revenue constraint:

(7) $$t \cdot \sum_{j=1}^{2} \gamma_j \left[\int_{\hat{w}_j}^{\bar{w}} awl(aw,t,\tau) dG_j(w) + \int_{\underline{w}}^{\hat{w}_j} wl(w,t,\tau) dG_j(w) \right] - \tau = 0,$$

assuming, for simplicity, that the government has no revenue needs.

Denote the optimal tax parameters by t^* and τ^*. Under some general conditions, one can show (see appendix A for details) that the optimal marginal tax rate is positive ($t^* > 0$), and the lump-sum tax is negative (τ^*<0). Thus, the labor income tax is progressive. Crucially note that as individuals differ in two characteristics (innate ability and scholastic aptitude), the non-negativity of the tax rate can not be ensured under the standard assumptions in the literature. To see the intuition for that, note that an individual with a given level of labor income can be of a relatively low innate ability but with high scholastic aptitude and vice versa. Thus, it may well be the case that income will be negatively correlated with well being.[5]

PT[5] **TP** Consider, for instance, two individuals: one whose innate ability is low but is gifted for schooling and the other with a reversed set of characteristics. If the income level of the former exceeds

Under such circumstances, a progressive labor income tax system would stand in conflict with maximization of an egalitarian social welfare function.

We turn next to address the question of the desirability of using an education tax as a supplement to the optimal linear labor income tax system. For this purpose, we denote an education tax (possibly, negative) by s and examine whether, starting from an optimal linear tax system and zero tax on education (s=0), levying a small tax on education would increase welfare.[6] Note that when an education tax is introduced the cost of acquiring education for type-j individuals is given by $e_j + s$. Denote the *Lagrangean* expression for the optimal tax problem by:

(8)

$$L = \sum_{j=1}^{2} \gamma_j \left[\int_{\hat{w}_j}^{\bar{w}} W[V(aw,t,\tau) - (e_j + s)] dG_j(w) + \int_{\underline{w}}^{\hat{w}_j} W[V(w,t,\tau)] dG_j(w) \right] +$$

$$\sum_{j=1}^{2} \lambda_j [V(a\hat{w}_j,t,\tau) - V(\hat{w}_j,t,\tau) - (e_j + s)] +$$

$$\mu \left\{ t \sum_{j=1}^{2} \gamma_j \left[\int_{\hat{w}_j}^{\bar{w}} awl(aw,t,\tau) dG_j(w) + \int_{\underline{w}}^{\hat{w}_j} wl(w,t,\tau) dG_j(w) \right] + s \cdot \sum_{j=1}^{2} \gamma_j [1 - G(\hat{w}_j)] - \tau \right\},$$

where λ_j, $j = 1,2$, and μ are *Lagrange* multipliers.

---

that of the latter, and the difference between the two income levels is sufficiently small, we obtain the negative correlation.

PT[6] **TP** Assuming that second order conditions are satisfied, this would imply that the optimal tax on education should be positive. Note that by education tax, we mean a (differential) lump-sum supplement to the labor income tax schedule. Thus, the education tax is paid over the whole working period of an individual, which is our stylized model is given by a single period.

We seek to sign the following derivative (employing the envelope theorem):

(9)

$$\frac{\partial L}{\partial s}\Big|_{s=0,t^*,\tau^*} = -\sum_{j=1}^{2}\gamma_j \int_{\hat{w}_j}^{\bar{w}} W'[V(aw,t^*,\tau^*)-e_j]dG_j(w) - \lambda_1 - \lambda_2 + \mu\left[\sum_{j=1}^{2}\gamma_j[1-G(\hat{w}_j)]\right]$$

$$= -Cov(\theta,K) - t^*/(1-t^*)\cdot E(\theta)\cdot\sum_{j=1}^{2}[\gamma_j \hat{w}_j G_j^{'}(\hat{w}_j)],$$

where $\theta \equiv W'$ denotes the social marginal utility of income, and $Cov(\cdot)$ and $E(\cdot)$ are the covariance and expectation operators; see appendix B for details.

The interpretation of equation (9) is as follows. The first term on the extreme right-hand side captures the re-distributive component. This term measures the direct redistributive effect of a unit increase in the tax on education, accompanied by an <u>optimal</u> adjustment in the (t^*,τ^*) pair. Note that the covariance term, $Cov(\theta,K)$, is typically negative, as individuals to whom the government assigns higher social marginal utility of income (and hence are those targeted by the government to obtain transfers) are more likely not to acquire education.[7] Thus, the first term works in the direction of levying a tax on education as an equity enhancing tool. The second term on the extreme right-hand side captures the disincentive component (the efficiency cost) of a tax on education. This disincentive effect is measured by the reduction in labor income tax revenues (and, correspondingly, the lump-sum transfer) due to the reduced incentive to acquire education, with a corresponding reduction in aggregate labor income (see appendix C for derivation). Note that the disincentive component is

[7] Heckman (2000) provides evidence of college participation by 18-24 year old high-school graduates in the US. He shows that from 1970-1993 the participation rate of the top half of the family income distribution has grown to over 75 percent; whereas, for the bottom 25 percent, participation has stagnated at the 40-45 percent range. Thus assuming that the covariance term in expression (9) is negatively signed seems plausible. However, one has to control for capital market imperfections (as we assume no market failures in our model). The extent to which credit issues induce under-representation of students from low socio-economic backgrounds in tertiary education is however debatable, as suggested by recent empirical evidence [see Carneiro and Heckman (2002)].

associated only with the reduced incentives to acquire education, as the tax on education does not distort the labor-leisure choice. This distortion exacerbates the already existing distortion associated with the positive labor income tax.[8] Thus, this term works in the opposite direction. Clearly, when the equity gains outweigh the efficiency costs, taxing education is desirable, and vice versa.

We turn next to provide sufficient conditions for the equity gains to indeed outweigh the efficiency costs, thereby establishing a case for an education tax. To facilitate our interpretation we re-formulate equation (9). Note that the direct effect of a unit increase in the tax on education on government revenues is given by the number of individuals who acquire education; namely, $\sum_{j=1}^{2}\gamma_j[1-G(\hat{w}_j)]$. Dividing equation (9) by the expression $\sum_{j=1}^{2}\gamma_j[1-G(\hat{w}_j)]$, and re-arranging terms, yields:

(9')

$$\frac{1}{\sum_{j=1}^{2}\gamma_j[1-G(\hat{w}_j)]}\cdot\frac{\partial L}{\partial s}\Big|_{s=0,t^*,\tau^*}=$$

$$-\frac{1}{\sum_{j=1}^{2}\gamma_j[1-G(\hat{w}_j)]}\cdot Cov(\theta,K)-t^*/(1-t^*)\cdot E(\theta)\cdot\sum_{j=1}^{2}\alpha_j\cdot\left[\frac{\hat{w}_j\cdot G_j^{'}(\hat{w}_j)}{1-G_j(\hat{w}_j)}\right],$$

where $\alpha_j \equiv \dfrac{\gamma_j[1-G(\hat{w}_j)]}{\sum_{j=1}^{2}\gamma_j[1-G(\hat{w}_j)]}$.

The latter equation describes the change in welfare per marginal dollar raised by the tax on education.

PT[8] **TP** Note that when the labor income tax rate is zero the second term vanishes.

Suppose that $W(\cdot)$ takes the common CES form; namely, $W(V) = V^{\rho} / \rho$, where $\rho < 1$ measures the degree of inequality aversion. Note that for the extreme case of a *Rawlsian* social planner, given by the limiting case where $\rho \to -\infty$, the covariance term on the right-hand side of equation (9') reduces to $\sum_{j=1}^{2} \gamma_j [1 - G(\hat{w}_j)]$, whereas the expectation term reduces to $E(\theta) = 1$. Thus, equation (9') reduces to:

(9'')

$$\frac{1}{\sum_{j=1}^{2} \gamma_j [1 - G(\hat{w}_j)]} \cdot \frac{\partial L}{\partial s}\Big|_{s=0,t^*,\tau^*} = 1 - t^* /(1 - t^*) \cdot \sum_{j=1}^{2} \alpha_j \cdot \left[\frac{\hat{w}_j \cdot G_j^{'}(\hat{w}_j)}{1 - G_j(\hat{w}_j)} \right].$$

Note that the second term on the right-hand side of equation (9'') captures the efficiency cost per marginal dollar raised by a unit increase in the education tax.

Following Diamond (1998), we assume that the term $w \cdot G_j^{'}(w) / [1 - G_j(w)]$, associated with the distribution of innate abilities for each type of individuals (j=1,2), is bounded from above.[9] This assumption is supported by empirical evidence. As shown by Saez (2001), analyzing US data, the term $wG'(w)/[1-G(w)]$ is increasing until very high productivity levels, above which the productivity distribution seems to be well approximated by a *Pareto* distribution. We denote by b_j the upper bound associated with the distribution $G_j(w)$, and let b denote the maximum between b_1 and b_2. It follows that the weighted average on the right-hand side of equation (9'')

PT[9] **TP**Diamond (1998) assumes that the distribution of innate abilities is single-peaked. He further assumes that for values above the modal skill level, innate abilities are distributed according to a *Pareto* distribution. It follows that the term $wG'(w)/[1-G(w)]$ is increasing up to the mode [as G'(w) is increasing up to the modal skill], and by virtue of the properties of the *Pareto* distribution [the density $G'(w)$ is proportional to $1/w^{(a+1)}$, for a>0], the term is constant above the modal skill level and given by a. Thus, the term $wG'(w)/[1-G(w)]$ is indeed bounded from above by the parameter a.

is bounded from above by the parameter *b*. Thus, given the labor income tax, the parameter *b* provides an upper bound on the distortion associated with an education tax (the induced reduction in labor income tax revenues). When the value of this parameter is sufficiently small, the disincentive effect is weak, and a case of taxing education is established.

An alternative interpretation of the *Rawlsian* case follows directly from equation (9), which, in such a case, reduces to the following expression:

$$(9''') \qquad \frac{\partial L}{\partial s}\Big|_{s=0,t^*,\tau^*} = \sum_{j=1}^{2} \gamma_j [1 - G(\hat{w}_j)] - t^* / (1 - t^*) \cdot \sum_{j=1}^{2} [\gamma_j \hat{w}_j G_j^{'}(\hat{w}_j)] \, .$$

The interpretation of condition is straightforward. In the *Rawlsian* case, as the most disadvantaged individual is assumed not to acquire education, a small tax on education affects her well-being only through the transfer, τ . Thus, a small tax on education is desirable, if and only if it gives rise to an increase in government revenues, thereby allowing the government to offer a higher lump-sum transfer. The first term on the right-hand side of equation (9''') measures the direct effect of a unit increase in the tax on education on government revenues. The second term on the right-hand side of equation (9''') measures the indirect effect of a tax on education; namely, the reduction in labor income tax revenues due to the reduced incentive to acquire education associated with a unit increase in the labor income tax. When the first (direct) effect of an increase in tax revenues exceeds the second (indirect) effect of reduction in tax revenues, a tax on education raises government revenues and, consequently, the lump-sum transfer financed by these revenues. A case for an education tax is thus established.

We state now the main result of the paper (for proof, see appendix D).

Proposition: For values of ρ and b sufficiently small, a small tax on education is welfare enhancing.[10]

The interpretation of the result is straightforward. When the objective of the government is sufficiently egalitarian (ρ small enough), and the distortion caused by an education tax per dollar raised is bounded (small *b*), a case for an education tax is established.

One particular interesting case to examine is when *h* is iso-elastic. One can show (see appendix E) that in the case of a *Rawlsian* social welfare function, a sufficient condition for the desirability of an education tax is that $\xi > b$, where ξ denotes the (constant) elasticity of labor supply. Thus, given the distribution of innate abilities, when labor-leisure distortion, as captured by the labor supply elasticity, is large enough, a case for an education tax is established. In such a case, the supplementary redistributive role of an education tax to the labor income tax is significant, due to the latter's large disincentive effect on labor-leisure choice. Note that unlike a labor income tax, which distorts both the labor-leisure choice and the education choice, an education tax distorts only the latter.

4. Conclusions

In this paper we have employed a simple model where individuals differ in both innate ability and scholastic aptitude and examined the desirability of taxing (subsidizing) education, as a supplement to an optimal labor income tax system. The rationale for taxing education derives from the informational constraint faced by the government who can not observe the individual characteristics themselves (and thus

PT[10] **TP**Provided that second order conditions are satisfied, it follows that at the optimum, the tax on education should be positive.

can not ideally base the tax on these characteristics). In a second best world, the variation in educational attainment across individuals can be used as a supplement to the observed variation in incomes, to infer about the individuals' unobserved characteristics, thereby enhancing the redistributive power of taxation.

It is important to qualify our results, by noting that there may well be other reasons to subsidize education; notably, the existence of an imperfect capital market which, given the uncertain nature of investment in human capital, imposes credit constraints on part of the population. In such a scenario, subsidizing education (directly or via tax breaks) is warranted on efficiency grounds, and it may well be the case, that all in all, the government would find it desirable to subsidize education.

References

Atkinson, A. and J. Stiglitz (1976) "The Design of Tax Structure: Direct versus Indirect Taxation", Journal of Public Economics, 6, 55-75

Barham, V., Boadway, R. and P. Pestieau (1995) "Education and the Poverty Trap", European Economic Review, 39, 1257-1275

Boadway, R., Marceau, N. and M. Marchand (1996) "Investment in Education and the Time Inconsistency of Redistributive Tax Policy", Economica, 63, 171-189

Boadway, R. and M. Marchand (1995) "The Use of Public Expenditure for Redistributive Purposes", Oxford Economic Papers, 47, 45-59

Brett, C. and J. Weymark (2003) "Financing Education Using Optimal Redistributive Taxation", Journal of Public Economics, 87, 2549-2569

Carneiro, P. and J. Heckman (2002) "The Evidence on Credit Constraints in Post-Secondary Schooling", The Economic Journal, 112, 705-734

Deaton, A. (1979) "Optimally Uniform Commodity Taxes", Economics Letters, 2, 357-361

Deaton, A. (1981) "Optimal Taxes and the Structure of Preferences", Econometrica, 49, 1245-1260

Diamond, P. (1998) "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates", American Economic Review, 88, 83-95

Eckstein, Z. and Y. Zilcha (1994) "The Effects of Compulsory Schooling on Growth, Income Distribution and Welfare", Journal of Public Economics, 54, 339-359

Gradstein, Mark. (2000) "An Economic Rationale for Public Education: The Value of Commitment", Journal of Monetary Economics, 45, 463-474

Heckman, J. (2000) "Policies to Foster Human Capital", Research in Economics, 54, 3-56

Mirrlees, J. (1971) "An Exploration in the Theory of Optimum Income Taxation", Review of Economic Studies, 38, 175-208

Razin, A. and E. Sadka (2001) "Labor, Capital and Finance: International Flows", Cambridge University Press

Saez, E. (2001) "Using Elasticities to Derive Optimal Tax Rates", Review of Economic Studies, 68, 205-229

Saint-Paul, G. (1994) "Unemployment, Wage Rigidity, and the Returns to Education", European Economic Review, 38, 535-543.

Sheshinski, E. (1971) "On the Theory of Optimal Income Taxation", Discussion Paper No. 172, Harvard Institute for Economic Research, Harvard University

Tuomala, M. (1986) "On the Optimal Income Taxation and Educational Decisions" Journal of Public Economics, 30, 183-198

Ulph, D. (1977) "On the Optimal Distribution of Income and Educational Expenditure", Journal of Public Economics, 8, 341-356

Wilson, A. (1999) "Education Policy: The Case for Government Intervention with Endogenous Abilities", Unpublished Manuscript, Department of Economics, Princeton University.

Appendix A: Proof of Non Negativity of the Marginal Tax Rate

We assume that second order conditions are satisfied, thus it suffices to show that the there exists a marginal welfare gain by slightly increasing the tax rate from t=0. Forming the *Lagrangean*, denoting by $\lambda_j, j = 1,2$ and μ, the multipliers associated with the incentive constraints in (5), for type j=1,2, and the revenue constraint in (7), respectively, we seek to sign the following expression (suppressing the tax parameters to abbreviate notation):

(A1)

$$\frac{\partial L}{\partial t}\bigg|_{t=0} = -\sum_{j=1}^{2}\gamma_j\left[\int_{\hat{w}_j}^{\bar{w}} aw\cdot l(aw)\cdot W'[V(aw)-e_j]dG_j(w) + \int_{\underline{w}}^{\hat{w}_j} w\cdot l(w)\cdot W'[V(w)]dG_j(w)\right]$$

$$+\sum_{j=1}^{2}\lambda_j\frac{\partial}{\partial t}[V(a\hat{w}_j)-V(\hat{w}_j)] + \mu\sum_{j=1}^{2}\gamma_j\left[\int_{\hat{w}_j}^{\bar{w}} awl(aw)dG_j(w) + \int_{\underline{w}}^{\hat{w}_j} wl(w)dG_j(w)\right]$$

By virtue of the social planner's optimization, one obtains the following two first-order conditions:

(A2) $$\frac{\partial L}{\partial \hat{w}_j}\Big|_{t=0} = \lambda_j\frac{\partial}{\partial \hat{w}_j}[V(a\hat{w}_j)-V(\hat{w}_j)] = 0 \Leftrightarrow \lambda_j = 0, j = 1,2\,;$$

(A3) $$\frac{\partial L}{\partial \tau}\Big|_{t=0} = \sum_{j=1}^{2}\gamma_j\left[\int_{\hat{w}_j}^{\bar{w}} W'[V(aw)-e_j]dG_j(w) + \int_{\underline{w}}^{\hat{w}_j} W'[V(w)]dG_j(w)\right] - \mu = 0$$

Substituting for $\lambda_j, j = 1,2$ and μ, from (A2) and (A3) into (A1), simplifying and re-arranging yields:

(A4) $$\frac{\partial L}{\partial t}\Big|_{t=0} = -Cov[\theta, I],$$

where $\theta \equiv W'$ denotes the social marginal utility of income, I denotes labor income and $Cov(\cdot)$ denotes the covariance operator.

Clearly, in general, one can not sign the right-hand side of equation (A4), for reasons discussed in the text. However, in two simple cases the sign would be unambiguously Positive: first, in the case of a *Rawlsian* planner, seeking to maximize the well being of the least well-off individual (necessarily the one with the lowest income level, as the individual with the lowest innate ability is assumed to acquire no schooling, and is therefore the one with the lowest income level); second, when the difference between scholastic aptitudes, $e_2 - e_1 \equiv \Delta$, $\Delta > 0$, is sufficiently small. The latter follows from the fact that when scholastic aptitude is identical for the two types ($\Delta = 0$), individuals with a higher level of income are necessarily those with higher innate ability, hence higher well being. By virtue of continuity, this extends to the case of small values of $\Delta > 0$.

More generally, when the correlation between innate ability and scholastic aptitude, and/or, the degree of inequality aversion exhibited by the social planner, is sufficiently high, the marginal tax rate will be positive.

Appendix B: Derivation of Equation (9)

For convenience we re-formulate equation (9), which is given by:

(B1)

$$\frac{\partial L}{\partial s}\Big|_{s=0,t^*,\tau^*} = -\sum_{j=1}^{2}\gamma_j\int_{\hat{w}_j}^{\bar{w}} W'[V(aw,t^*,\tau^*)-e_j]dG_j(w)-\lambda_1-\lambda_2+\mu\left[\sum_{j=1}^{2}\gamma_j[1-G(\hat{w}_j)]\right].$$

By virtue of the social planner's optimization (employing the envelope theorem for the individual choice problem), one obtains the two following first-order conditions:

(B2)

$$\frac{\partial L}{\partial \hat{w}_j}\Big|_{s=0,t^*,\tau^*} = \mu t^*\gamma_j\hat{w}_j[-al(a\hat{w}_j,t^*,\tau^*)+l(\hat{w}_j,t^*,\tau^*)]\cdot G_j'(\hat{w}_j)$$
$$+\lambda_j\frac{\partial}{\partial \hat{w}_j}[V(a\hat{w}_j,t^*,\tau^*)-V(\hat{w}_j,t^*,\tau^*)]=0;\ j=1,2,$$

(B3)

$$\frac{\partial L}{\partial \tau}\Bigg|_{s=0,t^*,\tau^*} = \sum_{j=1}^{2}\gamma_j\left[\int_{\hat{w}_j}^{\bar{w}} W'[V(aw,t^*,\tau^*)-e_j]dG_j(w)+\int_{\underline{w}}^{\hat{w}_j} W'[V(w,t^*,\tau^*)]dG_j(w)\right]-\mu=0$$

Substituting for λ_j, $j=1,2$ and μ, from (B2) and (B3) into (B1), employing (4), simplifying and re-arranging terms yields:

(B4) $$\frac{\partial L}{\partial s}\Big|_{s=0,t^*,\tau^*} = -Cov(\theta,K)-t^*/(1-t^*)\cdot E(\theta)\cdot\sum_{j=1}^{2}[\gamma_j\hat{w}_jG_j'(\hat{w}_j)].$$

Appendix C: The Indirect Effect of a Small Tax on Education

Denote by $R(s,t)$ the labor income tax revenues when the tax levied on education is s and the labor income tax rate is t. Formally, $R(s,t)$ is given by:

$$\text{(C1)} \qquad R(s,t) = t \cdot \sum_{j=1}^{2} \gamma_j \left[\int_{\hat{w}_j}^{\bar{w}} awl(aw,t) dG_j(w) + \int_{\underline{w}}^{\hat{w}_j} wl(w,t) dG_j(w) \right],$$

where $\hat{w}_j$, $j = 1,2$, is given by the implicit solution to:

$$\text{(C2)} \qquad V(a\hat{w}_j, t) - V(\hat{w}_j, t) - (e_j + s) = 0 .$$

The indirect effect of a small tax levied on education is formally given by the derivative $\frac{\partial R(s,t)}{\partial s}$.

Differentiation of the expression in (C1) yields:

$$\text{(C3)} \qquad \frac{\partial R(s,t)}{\partial s} = \left[t \cdot \sum_{j=1}^{2} \left[\gamma_j \hat{w}_j [-al(a\hat{w}_j, t) + l(\hat{w}_j, t)] \cdot G_j^{'}(\hat{w}_j) \right] \cdot \frac{\partial \hat{w}_j}{\partial s} \right].$$

Fully differentiating the expression in (C2) with respect to s and re-arranging terms yields:

$$\text{(C4)} \qquad \frac{\partial \hat{w}_j}{\partial s} = \frac{1}{(1-t) \cdot [al(a\hat{w}_j, t) - l(\hat{w}_j, t)]}.$$

Substituting from (C4) into (C3) yields:

$$\text{(C5)} \qquad \frac{\partial R(s,t)}{\partial s} = -t^* / (1 - t^*) \cdot \sum_{j=1}^{2} [\gamma_j \hat{w}_j G_j^{'}(\hat{w}_j)] .$$

The second term on the extreme right-hand side of equation (9) is given by the product of the expression on the right-hand side of (C5) and the social marginal utility of income [$E(\theta)$]. Thus, it measures in utility terms the reduction in labor income tax revenues associated with a unit increase in the education tax.

Appendix D: Proof of the Proposition

We prove the result for the limiting case where $\rho \to -\infty$ (a *Rawlsian* social planner). Then we extend the result by virtue of continuity.

To prove the desirability of an education tax, it is sufficient to show that the expression on the right-hand side of condition (9"), which is re-produced by condition (D1) for convenience, is positively signed:

(D1) $$\frac{1}{\sum_{j=1}^{2}\gamma_j[1-G(\hat{w}_j)]}\cdot\frac{\partial L}{\partial s}\Big|_{s=0,t^*,\tau^*} = 1-t^*/(1-t^*)\cdot\sum_{j=1}^{2}\alpha_j\cdot\left[\frac{\hat{w}_j\cdot G_j'(\hat{w}_j)}{1-G_j(\hat{w}_j)}\right],$$

where $\alpha_j \equiv \dfrac{\gamma_j[1-G(\hat{w}_j)]}{\sum_{j=1}^{2}\gamma_j[1-G(\hat{w}_j)]}$.

Employing the assumption that the term $w\cdot G_j'(w)/[1-G_j(w)]$ associated with the distribution of innate abilities for each type (j=1,2) is bounded from above, where the respective upper bounds are given by b_j, $j=1,2$, and further recalling that $b \equiv \max(b_1,b_2)$, it suffices to show that:[11]

(D2) $Sign[(1-t^*)/t^* - b] > 0$.

Differentiating the *Largrangean* with respect to *t*, taking the limit when $t \to 1$, we obtain:

(D3) $$\left(\frac{\partial L}{\partial t}\Big|_{t=1}\right) = \sum_{j=1}^{2}\gamma_j\left[\int_{\underline{w}}^{\hat{w}_j}\frac{\partial[w\cdot l(w)]}{\partial t}\Big|_{t=1}dG_j(w) + \int_{\hat{w}_j}^{\infty}\frac{\partial[aw\cdot l(aw)]}{\partial t}\Big|_{t=1}dG_j(w)\right].$$

PT[11] **TP**Recall that for the *Rawlsian* case, we show in appendix A that the marginal tax rate is positive.

The individual optimal choice of labor is given by the implicit solution to: $(1-t)w - h'(l) = 0$. Fully differentiating with respect to t yields: $\frac{\partial[w \cdot l(w)]}{\partial t} = \frac{-w^2}{h''[l(w)]} < 0$. Thus, it follows that:

(D4) $\left(\frac{\partial L}{\partial t}\Big|_{t=1} \right) < \frac{-\underline{w}^2}{h''(0)} < 0$.

As the upper bound on the derivative in (D4) is independent of the parameter *b*, it follows that for all *b* (including values arbitrarily close to zero) the optimal tax rate is bounded away from unity. Formally: $\lim_{b \to_+ 0} t^*(b) < 1$. By virtue of continuity, the result is proved for sufficiently small values of *b*. QED

Appendix E: A Sufficient Condition for the Desirability of an Education Tax for the Iso-elastic Case

As shown in appendix D [see equation (D2)], to prove that education tax is desirable, it suffices to show that:

(E1) $Sign[(1-t^*)/t^* - b] > 0$.

Note that the *Rawlsian* optimum tax rate is bounded from above by the *Laffer* rate, namely, the tax rate that maximizes overall tax revenues. With an iso-elastic functional form, that is, $h(l) = l^\alpha / \alpha$, with $\alpha \equiv 1/\xi + 1$, where ξ>0 denotes the (constant) labor supply elasticity, this would imply that:

(E2) $t^* < 1/(1+\xi) \Leftrightarrow (1-t^*)/t^* > \xi$,

To see this, denote by $R(t)$ the tax revenues when the tax rate is *t*, given by:

(E3) $$R(t) = t \cdot \sum_{j=1}^{2} \gamma_j \left[\int_{\hat{w}_j}^{\bar{w}} awl(aw,t) dG_j(w) + \int_{\underline{w}}^{\hat{w}_j} wl(w,t) dG_j(w) \right].$$

Assuming that second order conditions hold, it suffices to show that:

(E4) $$\left. \frac{\partial R(t)}{\partial t} \right|_{t = 1/(1+\xi)} < 0.$$

Differentiation of the expression on right-hand side of equation (E3) and re-arrangement yield:

(E5) $$\frac{\partial R(t)}{\partial t} = A(t) + B(t),$$

where:

$$A(t) \equiv \sum_{j=1}^{2} \gamma_j \left[\int_{\hat{w}_j}^{\bar{w}} \left[awl(aw,t) + t \cdot aw \frac{\partial l(aw,t)}{\partial t} \right] dG_j(w) + \int_{\underline{w}}^{\hat{w}_j} \left[wl(w,t) + t \cdot w \frac{\partial l(w,t)}{\partial t} \right] dG_j(w) \right],$$

and $$B(t) \equiv t \cdot \sum_{j=1}^{2} \gamma_j \left[\hat{w}_j \cdot \left[l(\hat{w}_j, t) - a \cdot l(a\hat{w}_j, t) \right] \cdot G_j^{'}(\hat{w}_j) \cdot \frac{\partial \hat{w}_j}{\partial t} \right].$$

We turn next to sign the two terms [$A(t)$ and $B(t)$] on the right-hand side of equation (E5), for $t = 1/(1+\xi)$.

One can show that the labor supply for the iso-elastic case is given by:

(E6) $\quad l(w,t) = [w(1-t)]^{\xi}$.

This implies that:

(E7) $\quad \dfrac{\partial l(w,t)}{\partial t} = -w\xi[w(1-t)]^{\xi-1}$.

Employing (E6) and (E7), it is straightforward to verify that, for any w:

(E8) $\quad \left[l(w,t) + t\dfrac{\partial l(w,t)}{\partial t} \right] \Bigg|_{t = 1/(1+\xi)} = 0.$

Substitution into the term $A(t)$ on the right-hand side of equation (E5), implies that $A[1/(1+\xi)] = 0$.

Fully differentiating equation (5) with respect to t and re-arranging terms imply that $\dfrac{\partial \hat{w}_j}{\partial t} > 0$. Moreover, it follows from (E6) that $\dfrac{\partial l(w,t)}{\partial w} > 0$. It is thus straightforward to verify, recalling that a>1, that the term $B(t)$ on the right-hand side of (E5) is negatively signed, for any positive value of t, hence $B[1/(1+\xi)] < 0$. Thus, we confirm that the derivative in (E4) is indeed negative hence the condition given in (E2) is confirmed.

Substitution from (E2) into (E1) implies that a sufficient condition for the desirability of an education tax is:

(E9) $\quad \xi > b$.

CESifo Working Paper Series

(for full list see www.cesifo.de)

1378 Theodore C. Bergstrom and John L. Hartman, Demographics and the Political Sustainability of Pay-as-you-go Social Security, January 2005

1379 Bruno S. Frey and Margit Osterloh, Yes, Managers Should Be Paid Like Bureaucrats, January 2005

1380 Oliver Hülsewig, Eric Mayer and Timo Wollmershäuser, Bank Loan Supply and Monetary Policy Transmission in Germany: An Assessment Based on Matching Impulse Responses, January 2005

1381 Alessandro Balestrino and Umberto Galmarini, On the Redistributive Properties of Presumptive Taxation, January 2005

1382 Christian Gollier, Optimal Illusions and Decisions under Risk, January 2005

1383 Daniel Mejía and Marc St-Pierre, Unequal Opportunities and Human Capital Formation, January 2005

1384 Luis H. R. Alvarez and Erkki Koskela, Optimal Harvesting under Resource Stock and Price Uncertainty, January 2005

1385 Ruslan Lukach, Peter M. Kort and Joseph Plasmans, Optimal R&D Investment Strategies with Quantity Competition under the Threat of Superior Entry, January 2005

1386 Alfred Greiner, Uwe Koeller and Willi Semmler, Testing Sustainability of German Fiscal Policy. Evidence for the Period 1960 – 2003, January 2005

1387 Gebhard Kirchgässner and Tobias Schulz, Expected Closeness or Mobilisation: Why Do Voters Go to the Polls? Empirical Results for Switzerland, 1981 – 1999, January 2005

1388 Emanuele Bacchiocchi and Alessandro Missale, Managing Debt Stability, January 2005

1389 Assar Lindbeck and Dirk Niepelt, Improving the SGP: Taxes and Delegation rather than Fines, January 2005

1390 James J. Heckman and Dimitriy V. Masterov, Skill Policies for Scotland, January 2005

1391 Emma Galli & Fabio Padovano, Sustainability and Determinants of Italian Public Deficits before and after Maastricht, January 2005

1392 Angel de la Fuente and Juan Francisco Jimeno, The Private and Fiscal Returns to Schooling and the Effect of Public Policies on Private Incentives to Invest in Education: A General Framework and Some Results for the EU, January 2005

1393 Juan C. Conesa and Carlos Garriga, Optimal Response to a Demographic Shock, January 2005

1394 Christian Gollier, Optimal Portfolio Management for Individual Pension Plans, February 2005

1395 Ruslan Lukach, Joseph Plasmans and Peter M. Kort, Innovation Strategies in a Competitive Dynamic Setting, February 2005

1396 Gebhard Kirchgässner, (Why) Are Economists Different?, February 2005

1397 Marko Köthenbürger, Panu Poutvaara and Paola Profeta, Why are More Redistributive Social Security Systems Smaller? A Median Voter Approach, February 2005

1398 Gabrielle Demange, Free Choice of Unfunded Systems: A First Assessment, February 2005

1399 Carlos Fonseca Marinheiro, Sustainability of Portuguese Fiscal Policy in Historical Perspective, February 2005

1400 Roel M. W. J. Beetsma and Koen Vermeylen, The Effect of Monetary Unification on Public Debt and its Real Return, February 2005

1401 Frank Asche, Petter Osmundsen and Maria Sandsmark, Is It All Oil?, February 2005

1402 Giacomo Corneo, Media Capture in a Democracy: The Role of Wealth Concentration, February 2005

1403 A. Lans Bovenberg and Thijs Knaap, Ageing, Funded Pensions and the Dutch Economy, February 2005

1404 Thiess Büttner, The Incentive Effect of Fiscal Equalization Transfers on Tax Policy, February 2005

1405 Luisa Fuster, Ayşe İmrohoroğlu and Selahattin İmrohoroğlu, Personal Security Accounts and Mandatory Annuitization in a Dynastic Framework, February 2005

1406 Peter Claeys, Policy Mix and Debt Sustainability: Evidence from Fiscal Policy Rules, February 2005

1407 James M. Malcomson, Supplier Discretion over Provision: Theory and an Application to Medical Care, February 2005

1408 Thorvaldur Gylfason, Interview with Assar Lindbeck, February 2005

1409 Christian Gollier, Some Aspects of the Economics of Catastrophe Risk Insurance, February 2005

1410 Gebhard Kirchgässner, The Weak Rationality Principle in Economics, February 2005

1411 Carlos José Fonseca Marinheiro, Has the Stability and Growth Pact Stabilised? Evidence from a Panel of 12 European Countries and Some Implications for the Reform of the Pact, February 2005

1412 Petter Osmundsen, Frank Asche, Bård Misund and Klaus Mohn, Valuation of International Oil Companies –The RoACE Era, February 2005

1413 Gil S. Epstein and Shmuel Nitzan, Lobbying and Compromise, February 2005

1414 Marcel F. M. Canoy, Jan C. van Ours and Frederick van der Ploeg, The Economics of Books, February 2005

1415 Eric A. Hanushek and Ludger Wößmann, Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries, February 2005

1416 George Kapetanios and M. Hashem Pesaran, Alternative Approaches to Estimation and Inference in Large Multifactor Panels: Small Sample Results with an Application to Modelling of Asset Returns, February 2005

1417 Samuel Mühlemann, Jürg Schweri, Rainer Winkelmann and Stefan C. Wolter, A Structural Model of Demand for Apprentices. February 2005

1418 Giorgio Brunello and Lorenzo Rocco, Educational Standards in Private and Public Schools, February 2005

1419 Alex Bryson, Lorenzo Cappellari and Claudio Lucifora, Why so Unhappy? The Effects of Unionisation on Job Satisfaction, March 2005

1420 Annalisa Luporini, Relative Performance Evaluation in a Multi-Plant Firm, March 2005

1421 Giorgio Bellettini and Carlotta Berti Ceroni, When the Union Hurts the Workers: A Positive Analysis of Immigration Policy, March 2005

1422 Pieter Gautier, Michael Svarer and Coen Teulings, Marriage and the City, March 2005

1423 Ingrid Ott and Stephen J. Turnovsky, Excludable and Non-Excludable Public Inputs: Consequences for Economic Growth, March 2005

1424 Frederick van der Ploeg, Back to Keynes?, March 2005

1425 Stephane Dees, Filippo di Mauro, M. Hashem Pesaran and L. Vanessa Smith, Exploring the International Linkages of the Euro Area: a Global VAR Analysis, March 2005

1426 Hans Pitlik, Friedrich Schneider and Harald Strotmann, Legislative Malapportionment and the Politicization of Germany's Intergovernmental Transfer System, March 2005

1427 Konstantinos Angelopoulos and Apostolis Philippopoulos, The Role of Government in Anti-Social Redistributive Activities, March 2005

1428 Ansgar Belke and Daniel Gros, Asymmetries in the Trans-Atlantic Monetary Policy Relationship: Does the ECB follow the Fed?, March 2005

1429 Sören Blomquist and Luca Micheletto, Optimal Redistributive Taxation when Government's and Agents' Preferences Differ, March 2005

1430 Olof Åslund and Peter Fredriksson, Ethnic Enclaves and Welfare Cultures – Quasi-Experimental Evidence, March 2005

1431 Paul De Grauwe, Roberto Dieci and Marianna Grimaldi, Fundamental and Non-Fundamental Equilibria in the Foreign Exchange Market. A Behavioural Finance Framework, March 2005

1432 Peter Egger, Stefan Gruber, Mario Larch and Michael Pfaffermayr, Knowledge-Capital Meets New Economic Geography, March 2005

1433 George Economides and Apostolis Philippopoulos, Should Green Governments Give Priority to Environmental Policies over Growth-Enhancing Policies?, March 2005

1434 George W. Evans and Seppo Honkapohja, An Interview with Thomas J. Sargent, March 2005

1435 Helge Berger and Volker Nitsch, Zooming Out: The Trade Effect of the Euro in Historical Perspective, March 2005

1436 Marc-Andreas Muendler, Rational Information Choice in Financial Market Equilibrium, March 2005

1437 Martin Kolmar and Volker Meier, Intra-Generational Externalities and Inter-Generational Transfers, March 2005

1438 M. Hashem Pesaran and Takashi Yamagata, Testing Slope Homogeneity in Large Panels, March 2005

1439 Gjermund Nese and Odd Rune Straume, Industry Concentration and Strategic Trade Policy in Successive Oligopoly, April 2005

1440 Tomer Blumkin and Efraim Sadka, A Case for Taxing Education, April 2005