UNIVERSITY of York

Discussion Papers in Economics

No. 18/07

On Public Education Spending under Nonlinear Income Taxation

Alan Krause

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

On Public Education Spending under Nonlinear Income Taxation

Alan Krause*
University of York

20 August 2018

Abstract

This paper examines a model in which public education spending is skill specific. It may be directed towards low-skill or high-skill individuals, increasing their respective skills and wages. Education spending is financed by nonlinear income taxation. We show that the tax and education-spending policy most preferred by low-skill individuals may include more education spending for high-skill individuals than for themselves. The tax and education-spending policy most preferred by high-skill individuals has no spending on education for low-skill individuals. Our results provide support for previous findings that education policy should favour the high-skilled, despite the government's redistributive goals.

Keywords: public education; nonlinear income taxation.

JEL Classifications: H21, H42.

*Department of Economics and Related Studies, University of York, Heslington, York, YO10 5DD, U.K. E-mail: alan.krause@york.ac.uk.

1 Introduction

Previous research has shown that it can be socially optimal for the government to set policy (such as education policy) to increase skill/wage differences in Mirrlees (1971) style nonlinear income tax models. See, for example, Krause (2006), Cremer, et al. (2011), and Stephens (2012). The result is of theoretical interest, because it shows that increasing skill differences can be socially desirable, even though it is such differences that necessitate second-best (distortionary) redistributive taxation in the first place. The result may also be of policy interest, because it provides a justification for favouring education spending which benefits high-skill individuals (e.g., subsidising universities) over that which benefits low-skill individuals (e.g., vocational training) despite the government's redistributive goals. In the present paper, we show that this result is surprisingly robust in the following respect. An examination of individual (rather than social) preferences reveals that the tax and education-spending policy most preferred by low-skill individuals may include *more* spending on education for high-skill individuals than on education for themselves. The tax and education-spending policy most preferred by high-skill individuals includes education spending only for themselves; they prefer that there be no spending on education for low-skill individuals. These findings hold even though the only effect education has in our model is to increase wages: low-skill education increases the low-skill type's wage, and high-skill education increases the high-skill type's wage. The cost of acquiring education is assumed to be the same for both types, and is set equal to zero.[1]

The intuition for our findings can be summarised as follows. The tax and education-spending policy most preferred by low-skill individuals corresponds to that chosen under a maxi-min social welfare objective. Spending on education for low-skill individuals is desirable, since it is their welfare that is to be maximised. However, spending on education for high-skill individuals is also desirable, to the extent that it facilitates redistribution from high-skill to low-skill individuals. Indeed, redistribution under maxi-

[1] It is sometimes assumed in the literature that the low-skill type's effort cost of acquiring education is greater than that of the high-skill type. Such an assumption in our model would strengthen our main findings.

min is extreme, which is why this motive for spending on high-skill education may outweigh the motive for spending on low-skill education. On the other hand, the tax and education-spending policy most preferred by high-skill individuals corresponds to that chosen under a maxi-max social welfare objective. In this case, policy is set to redistribute from low-skill to high-skill individuals. However, spending on education for low-skill individuals does not facilitate this redistribution. Thus, there is no spending on low-skill education, while spending on high-skill education is desirable because it is the welfare of high-skill individuals that is to be maximised.

Our paper is also related to the literature on voting over nonlinear income tax schedules, although this literature does not include education spending. See, for example, Roell (2012) and Brett and Weymark (2016, 2017). Rather than choosing the nonlinear income tax schedule that maximises a utilitarian social welfare function, this literature derives the nonlinear income tax schedule most preferred by each type of individual. For the lowest-skill (resp. highest-skill) individuals, this will be the tax schedule chosen under a maxi-min (resp. maxi-max) social welfare objective. Under certain conditions (single-peaked preferences), the nonlinear income tax schedule chosen under the majority decision rule will be that most preferred by middle-skill individuals (median voter theorem).

The literature on tax and education policy is extensive, and it deals with many aspects of education not considered here, such as: private expenditure on education, dynamic issues, and whether education expenditure should be taxed or subsidised. A sample of this literature includes Ulph (1977), Hare and Ulph (1979), Grout (1983), Tuomala (1986), Boadway, et al. (1996), De Fraja (2002), Fleurbaey, et al. (2002), Brett and Weymark (2003), Bovenberg and Jacobs (2005), Blumkin and Sadka (2008), Maldonado (2008), Krause (2009), and Guo and Krause (2013). Compared to this literature, we consider only the most salient feature of education, i.e., its role in increasing skills and wages. We then show that this feature alone is sufficient to justify an education policy which favours high-skill individuals. Specifically, we show that such a policy is broadly consistent with the preferences of both low-skill and high-skill individuals.

The remainder of the paper is organised as follows. Section 2 presents our analytical framework, Section 3 presents and discusses our results, and Section 4 concludes. Proofs are contained in an appendix.

2 Model and Assumptions

The model is a standard two-type version of the Mirrlees (1971) model, but extended to include skill-specific public education spending. There is a unit measure of individuals, with a proportion $\phi \in (0,1)$ being high-skill workers and $(1-\phi)$ being low-skill workers. The high-skill type's wage is denoted by $w_H = \omega_H(1+f(e_H))$, where ω_H can be interpreted as the high-skill type's endowment skill level or wage rate, and e_H is public spending on high-skill education. It is assumed that $f(0)=0$, $f'(e_H)>0$, and $f''(e_H)<0$. Therefore, if there is no spending on high-skill education, the high-skill type's wage is equal to their endowment wage. Analogously, the low-skill type's wage is denoted by $w_L = \omega_L(1+f(e_L))$. It can be seen that the only effect education spending has in our model is to increase wages. We assume throughout the paper that $w_H > w_L$ for all relevant values of e_H and e_L.

Both types have the same preferences, represented by the quasi-linear utility function $x_i - v(l_i)$, where x_i is type i's consumption and l_i is type i's labour supply.[2] The function $v(l_i)$ is assumed to be increasing and strictly convex. Let $y_i = w_i l_i$ denote type i's pre-tax income; thus $y_i - x_i$ is the income tax paid by a type i individual.

Our interest is in the tax and education-spending policy most preferred by each type of individual. This can be found by solving the following problem. Choose a tax treatment for the low-skill individuals, $\langle y_L, x_L \rangle$, and for the high-skill individuals, $\langle y_H, x_H \rangle$, as well as spending on low-skill and high-skill education, e_L and e_H, to maximise:

$$x_i - v\left(\frac{y_i}{w_i}\right) \qquad i = L \text{ or } H \tag{2.1}$$

subject to:

$$(1-\phi)(y_L - x_L) + \phi(y_H - x_H) \geq e_L + e_H \tag{2.2}$$

[2] The implications of relaxing the quasi-linear utility specification are discussed in Subsection 3.1.

$$x_L - v\left(\frac{y_L}{w_L}\right) \geq x_H - v\left(\frac{y_H}{w_L}\right) \tag{2.3}$$

$$x_H - v\left(\frac{y_H}{w_H}\right) \geq x_L - v\left(\frac{y_L}{w_H}\right) \tag{2.4}$$

$$x_i \geq \overline{x} \qquad\qquad i = L \textit{ and } H \tag{2.5}$$

When $i = L$ (resp. $i = H$) in the objective function (2.1), the solution to programme (2.1) – (2.5) represents the low-skill (resp. high-skill) type's most preferred tax and education-spending policy. This is the tax and education-spending policy that each type would choose to maximise their own welfare. Equation (2.2) is the budget constraint, requiring that total income tax revenues be greater than or equal to total spending on education. Equations (2.3) and (2.4) are incentive-compatibility (IC) constraints. The key feature of Mirrlees-style models is that each individual's skill type is unobservable, so the tax treatments must satisfy the IC constraints. Satisfaction of the IC constraints ensures that the tax treatments are monotonic, $\langle y_H, x_H\rangle \geq \langle y_L, x_L\rangle$.[3] When $i = L$ (maxi-min) constraint (2.4) will be binding and constraint (2.3) will be slack. This is because the objective is to choose the best tax and education-spending policy for low-skill individuals, which creates an incentive for high-skill individuals to mimic low-skill individuals, but not vice versa. The reverse is true when $i = H$ (maxi-max); constraint (2.3) will bind and constraint (2.4) will be slack. Finally, equation (2.5) requires that each type receive at least a survival level of consumption, $\overline{x} > 0$. This constraint is similar to the minimum-utility constraints in Roell (2012) and Brett and Weymark (2016). We assume that constraint (2.5) is slack under maxi-min, because the objective is to maximise the welfare of low-skill individuals, which will include a sufficient level of low-skill consumption. Monotonicity then ensures that the high-skill type's consumption level is also sufficient. However, constraint (2.5) may bind under maxi-max, as the chosen tax and education-spending policy will seek to provide low-skill individuals with minimal (zero) consumption.

[3] Monotonicity is a standard implication of incentive compatibility. See, for example, chapter 3 in Laffont and Martimort (2002).

3 Results

Our main results are presented in two propositions (proofs are in the appendix):

Proposition 1 *The tax and education-spending policy most preferred by low-skill individuals will include more spending on high-skill education than on low-skill education if and only if the following condition is satisfied*:

$$v'(l_L)l_L < \phi\left[v'(l_H)l_H - v'(l_M)l_M\right] \tag{3.1}$$

Equation (3.1) presents the condition under which low-skill individuals prefer that $e_H > e_L$, and it can be interpreted as follows. The left-hand side of equation (3.1) reflects the labour supply of low-skill individuals under their preferred tax and education-spending policy. Naturally, low-skill individuals will want to choose a policy that minimises their labour supply; this acts towards satisfying the inequality in equation (3.1). As low-skill individuals benefit from education via increased wages, and their benefit from increased wages is proportional to their labour supply, a low level of low-skill labour supply is associated with a lower preference for low-skill education.

The right-hand side of equation (3.1) reflects the difference between the labour supply of high-skill individuals and of high-skill individuals who mimic (denoted by M) under the low-skill type's most preferred policy. Low-skill individuals will want high-skill individuals to work a lot, generating income available for redistribution. Therefore, they want l_H to be large. A high-skill individual who mimics chooses to work less and earn the pre-tax income intended for low-skill individuals, which is y_L. Since $y_L = w_L l_L$ and low-skill individuals want to minimise their own labour supply, their preferred level of y_L will also be minimal. And since $l_M = y_L/w_H$, the mimicker's labour supply l_M will also be minimal. These forces together act towards increasing the right-hand side of equation (3.1), and hence towards satisfying the condition. As high-skill education raises the high-skill type's wage, it increases their opportunity cost of choosing to work less. Therefore, an increase in the right-hand side of equation (3.1) is associated with low-skill individuals having a stronger preference for high-skill education, as it helps deter mimicking.

The preceding discussion of Proposition 1 suggests that low-skill individuals will prefer a tax and education-spending policy that satisfies equation (3.1), when possible. We now present numerical examples using plausible parameter values in which equation (3.1) is satisfied and violated. In the numerical examples, the utility function takes the form:

$$x_i - \frac{(l_i)^{1+\frac{1}{\eta}}}{1+\frac{1}{\eta}} \tag{3.2}$$

where $\eta > 0$ is the labour supply elasticity. The wage equation takes the form $w_i = \omega_i(1 + e_i^{\alpha})$, where $\alpha \in (0, 1)$. According to the OECD (2017), approximately one-third of individuals have tertiary-level education. We therefore assume that one-third of the population is high-skill and the remainder are low-skill, i.e., $\phi = 1/3$. Estimates of the college wage premium are around 60% (see Fang 2006 and Goldin and Katz 2007). We therefore normalise the low-skill type's endowment wage to unity ($\omega_L = 1$), and set the high-skill type's endowment wage to 1.6 ($\omega_H = 1.6$). Finally, we arbitrarily set $\alpha = 0.25$, and we consider three plausible values of the labour supply elasticity: $\eta = 1/3$, $\eta = 1/2$, and $\eta = 1$. The results are shown in Table 1. When $\eta = 1/3$, condition (3.1) is violated and low-skill individuals prefer $e_L > e_H$. But when $\eta = 1/2$ or $\eta = 1$, condition (3.1) is satisfied and low-skill individuals prefer $e_H > e_L$. Empirical estimates of the labour supply elasticity suggest that a value of $0.5 - 1$ is quite plausible, especially when based on aggregate data (Chetty, et al. 2011).

Our second proposition states the key feature of the tax and education-spending policy most preferred by high-skill individuals:

Proposition 2 *The tax and education-spending policy most preferred by high-skill individuals has zero spending on education for low-skill individuals.*

The only possible motive that high-skill individuals could have for preferring some spending on low-skill education would be if it facilitated redistribution from low-skill to high-skill individuals. However, a low-skill individual who mimics works longer (to earn y_H) than low-skill individuals who reveal themselves (and earn y_L). This is the exact opposite of mimicking behaviour under the low-skill type's most preferred policy (cf. the right-hand side of equation (3.1) above). Indeed, if it were possible, high-skill individuals would prefer negative spending on low-skill education to reduce the low-skill type's wage, in order to deter mimicking.[4] However, negative education spending is not possible, so the best that high-skill individuals can do is choose zero spending on low-skill education.

3.1 Discussion: Other Utility Functions

When the utility function is quasi-linear in consumption, as assumed above, an analytical condition such as equation (3.1) is easily derived. This is because the Lagrange multipliers can be solved as functions of the model's parameters. This is not the case when the utility function is quasi-linear in labour, $u(x_i) - l_i$, with $u(x_i)$ increasing and strictly concave, or when it takes the more general form $u(x_i) - v(l_i)$. However, one might wonder if our results change under more general utility specifications. The short answer is that our main findings remain intact.

Suppose the utility function takes the form:

$$\frac{(x_i)^{1-\sigma}}{1-\sigma} - \frac{(l_i)^{1+\frac{1}{\eta}}}{1+\frac{1}{\eta}} \tag{3.3}$$

where $\sigma > 0$ is the coefficient of relative risk aversion. When $\sigma = 1$ the utility function becomes logarithmic in consumption. Tables 2 and 3 are analogous to Table 1, except they use the more general utility specification. Table 2 assumes that $\sigma = 1$ (log utility), and it shows that low-skill individuals prefer some spending on high-skill education, though this is less than they prefer for themselves. However, an example in which low-skill individuals prefer $e_H > e_L$ is easily obtained. This is shown in Table 3, which assumes that $\sigma = 0.1$. Low-skill individuals prefer $e_H > e_L$ when the labour supply elasticity is sufficiently high.

Regarding the high-skill type's most preferred tax and education-spending policy, it still has $e_L = 0$, so high-skill individuals still prefer that all education spending be for themselves. The intuition is the same as that underlying Proposition 2, discussed above.

[4] This is similar to the finding in Krause (2006) that vocational training should be 'under-provided' relative to the Samuelson rule along segments of the second-best Paretian frontier in which the low-skill type's IC constraint is binding.

4 Concluding Comments

Public spending on education, beyond the primary and secondary levels, can be broadly classified as spending on universities or on vocational training programmes. It is generally true that high-ability individuals attend university, which has raised concern that public subsidisation of universities is inconsistent with society's redistributive objectives. Much of the previous literature on tax and education policy has concluded that these concerns are unwarranted. In the present paper, we provide support for this conclusion by considering preferences at the individual (rather than social) level. In particular, we show it is possible that the tax and education-spending policy most preferred by low-skill individuals includes more spending on high-skill education than on low-skill education.

5 Appendix

Proof of Proposition 1

When $i = L$ in the objective function (2.1), the first-order conditions on x_L, x_H, e_L, and e_H in programme (2.1) – (2.5) can be written as, respectively:

$$1 - \lambda(1-\phi) - \theta_H = 0 \tag{A.1}$$

$$-\lambda\phi + \theta_H = 0 \tag{A.2}$$

$$\frac{w_L'(e_L)}{w_L^2} v'\left(\frac{y_L}{w_L}\right) y_L - \lambda = 0 \tag{A.3}$$

$$\frac{\theta_H w_H'(e_H)}{w_H^2}\left[v'\left(\frac{y_H}{w_H}\right) y_H - v'\left(\frac{y_L}{w_H}\right) y_L\right] - \lambda = 0 \tag{A.4}$$

where $\lambda > 0$ and $\theta_H > 0$ are the multipliers on constraints (2.2) and (2.4), respectively. Solving (A.1) and (A.2) for the multipliers yields $\lambda = 1$ and $\theta_H = \phi$. Equations (A.3) and (A.4) can be rewritten and rearranged to yield:

$$\frac{w_L'(e_L)}{w_L} = \frac{1}{v'(l_L)\, l_L} \tag{A.5}$$

$$\frac{w_H'(e_H)}{w_H} = \frac{1}{\phi\left[v'\left(l_H\right)l_H - v'\left(l_M\right)l_M\right]} \tag{A.6}$$

From the wage equation $w_i = \omega_i(1 + f(e_i))$ we obtain:

$$\frac{w_i'(e_i)}{w_i} = \frac{f'(e_i)}{1 + f(e_i)} \tag{A.7}$$

and:

$$\frac{\partial \frac{w_i'(e_i)}{w_i}}{\partial e_i} = \frac{f''(e_i)(1 + f(e_i)) - f'(e_i)f'(e_i)}{(1 + f(e_i))^2} < 0 \tag{A.8}$$

which implies that:

$$e_H > e_L \qquad \Longleftrightarrow \qquad \frac{w_H'(e_H)}{w_H} < \frac{w_L'(e_L)}{w_L} \tag{A.9}$$

Using equation (A.9), manipulation of (A.5) and (A.6) yields equation (3.1). ■

Proof of Proposition 2

To prove that high-skill individuals prefer $e_L = 0$, suppose to the contrary that $e_L > 0$ in the solution to programme (2.1) – (2.5) when $i = H$ in the objective function (2.1). Then regardless of whether constraint (2.5) is or is not binding, the following first-order condition on e_L must hold:

$$\frac{w_L'(e_L)}{w_L^2}\left[v'\left(\frac{y_L}{w_L}\right)y_L - v'\left(\frac{y_H}{w_L}\right)y_H\right] = \frac{\lambda}{\theta_L} > 0 \tag{A.10}$$

where $\lambda > 0$ and $\theta_L > 0$ are the multipliers on constraints (2.2) and (2.3), respectively. But equation (A.10) cannot be satisfied, because $y_H \geq y_L$ by monotonicity. ■

References

[1] Blumkin, T. and E. Sadka (2008), "A Case for Taxing Education", *International Tax and Public Finance*, 15, 145-163.

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

[3] Bovenberg, A. and B. Jacobs (2005), "Redistribution and Education Subsidies are Siamese Twins", *Journal of Public Economics*, 89, 2005-2035.

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

[5] Brett, C. and J. Weymark (2016), "Voting over Selfishly Optimal Nonlinear Income Tax Schedules with a Minimum-Utility Constraint", *Journal of Mathematical Economics*, 67, 18-31.

[6] Brett, C. and J. Weymark (2017), "Voting over Selfishly Optimal Nonlinear Income Tax Schedules", *Games and Economic Behavior*, 101, 172-188.

[7] Chetty, R., A. Guren, D. Manoli and A. Weber (2011), "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins", *American Economic Review*, 101, 471-475.

[8] Cremer, H., P. Pestieau and M. Racionero (2011), "Unequal Wages for Equal Utilities", *International Tax and Public Finance*, 18, 383-398.

[9] De Fraja, G. (2002), "The Design of Optimal Education Policies", *Review of Economic Studies*, 69, 437-466.

[10] Fang, H. (2006), "Disentangling the College Wage Premium: Estimating a Model with Endogenous Education Choices", *International Economic Review*, 47, 1151-1185.

[11] Fleurbaey, M., R. Gary-Bobo and D. Maguain (2002), "Education, Distributive Justice, and Adverse Selection", *Journal of Public Economics*, 84, 113-150.

[12] Goldin, C. and L. Katz (2007), "The Race between Education and Technology: The Evolution of U.S. Educational Wage Differentials, 1890-2005", *NBER Working Paper*, No. 12984.

[13] Grout, P. (1983), "Imperfect Information, Markets and Public Provision of Education", *Journal of Public Economics*, 22, 113-121.

[14] Guo, J-T. and A. Krause (2013), "Optimal Nonlinear Taxation of Income and Education Expenditures", *Oxford Economic Papers*, 65, 74-95.

[15] Hare, P. and D. Ulph (1979), "On Education and Distribution", *Journal of Political Economy*, 87, S193-S212.

[16] Krause, A. (2006), “Redistributive Taxation and Public Education”, *Journal of Public Economic Theory*, 8, 807-819.

[17] Krause, A. (2009), “Education and Taxation Policies in the Presence of Countervailing Incentives”, *Economica*, 76, 387-399.

[18] Laffont, J. and D. Martimort (2002), *The Theory of Incentives*: *The Principal-Agent Model*, Princeton University Press.

[19] Maldonado, D. (2008), “Education Policies and Optimal Taxation”, *International Tax and Public Finance*, 15, 131-143.

[20] Mirrlees, J. (1971), “An Exploration in the Theory of Optimum Income Taxation”, *Review of Economic Studies*, 38, 175-208.

[21] OECD (2017), *Education at a Glance*, OECD publications.

[22] Roell, A. (2012), “Voting over Nonlinear Income Tax Schedules”, *mimeo.*

[23] Stephens, E. (2012), “Teach a Man to Fish? Education vs. Optimal Taxation”, *Canadian Journal of Economics*, 45, 1700-1727.

[24] Tuomala, M. (1986), “On the Optimal Income Taxation and Educational Decisions”, *Journal of Public Economics*, 30, 183-198.

[25] Ulph, D. (1977), “On the Optimal Distribution of Income and Educational Expenditure”, *Journal of Public Economics*, 8, 341-356.

TABLE 1

Low-Skill Type's Most Preferred Tax and Education-Spending Policy

Utility Function: $x_i - \frac{(l_i)^{1+\frac{1}{\eta}}}{1+\frac{1}{\eta}}$

Parameter values			
ϕ	0.33	ω_L	1.00
α	0.25	ω_H	1.60
	$\eta = 1/3$	$\eta = 1/2$	$\eta = 1$
e_L	0.103	0.112	0.141
e_H	0.095	0.114	0.233
w_L	1.566	1.579	1.613
w_H	2.488	2.530	2.712
x_L	1.739	1.828	2.250
y_L	1.618	1.690	1.966
x_H	2.537	3.070	5.665
y_H	3.372	4.024	7.355

TABLE 2

Low-Skill Type's Most Preferred Tax and Education-Spending Policy

Utility Function: $\ln(x_i) - \frac{(l_i)^{1+\frac{1}{\eta}}}{1+\frac{1}{\eta}}$

	Parameter values		
ϕ	0.33	ω_L	1.00
α	0.25	ω_H	1.60
σ	1.00		
	$\eta = ⅓$	$\eta = ½$	$\eta = 1$
e_L	0.085	0.087	0.092
e_H	0.067	0.064	0.057
w_L	1.540	1.542	1.550
w_H	2.415	2.405	2.381
x_L	1.442	1.386	1.303
y_L	1.355	1.312	1.255
x_H	1.958	2.023	2.125
y_H	2.590	2.622	2.668

TABLE 3

Low-Skill Type's Most Preferred Tax and Education-Spending Policy

Utility Function: $\frac{(x_i)^{1-\sigma}}{1-\sigma} - \frac{(l_i)^{1+\frac{1}{\eta}}}{1+\frac{1}{\eta}}$

Parameter values			
ϕ	0.33	ω_L	1.00
α	0.25	ω_H	1.60
σ	0.10		
	$\eta = 1/3$	$\eta = 1/2$	$\eta = 1$
e_L	0.100	0.108	0.126
e_H	0.091	0.105	0.180
w_L	1.563	1.573	1.596
w_H	2.478	2.511	2.642
x_L	1.695	1.750	1.990
y_L	1.579	1.623	1.768
x_H	2.450	2.881	4.625
y_H	3.255	3.773	5.987