DISCUSSION PAPER SERIES

No. 4201

INTERGENERATIONAL TRANSFER OF HUMAN CAPITAL AND OPTIMAL EDUCATION POLICY

Helmuth Cremer and Pierre Pestieau

PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP4201.asp**

ISSN 0265-8003

INTERGENERATIONAL TRANSFER OF HUMAN CAPITAL AND OPTIMAL EDUCATION POLICY

Helmuth Cremer, Université des Sciences Sociales de Toulouse and CEPR
Pierre Pestieau, Université de Liège and CEPR

Discussion Paper No. 4201
January 2004

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **PUBLIC POLICY**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Helmuth Cremer and Pierre Pestieau

CEPR Discussion Paper No. 4201

January 2004

ABSTRACT

Intergenerational Transfer of Human Capital and Optimal Education Policy*

This Paper studies the design of education policies in a setting of successive generations with heterogeneous individuals (high and low earning ability). Parents' investment in education is motivated by warm-glow altruism and determines the probability that a child has high ability. Education policies consist of a subsidy on private educational investments and possibly of public education. We show that when an income tax is available, the subsidy on education should not depend on redistributive considerations. Instead, it is determined by two terms. First, a Pigouvian term that arises because under warm-glow altruism parents' utility does not properly account for the impact of education on future generations. The second term captures a 'merit good' effect, which arises when the warm-glow term is not fully included in social welfare (possibility of laundering out). The two terms are of opposite sign and the optimal subsidy may be positive or negative. Finally, we derive conditions under which public education is welfare-improving and show that total crowding out of private expenditure (for one of the types) may be desirable.

JEL Classification: H21, H23, I20 and I28
Keywords: education policy, human capital, intergenerational transfers and redistribution

Helmuth Cremer
Institut d'Economie Industrielle
Université des Sciences Sociales
21 Allée de Brienne
F-31000 Toulouse
FRANCE
Tel: (33 5) 6112 8606
Fax: (33 5) 6112 8637
Email: helmut@cict.fr

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=124993

Pierre Pestieau
Centre de Recherche en Economie Publique
et de la Population (CREPP)
Université de Liège
Boulevard du Rectorat, 7 (B31)
B-4000 Liège
BELGIUM
Tel: (32 4) 366 3108
Fax: (32 4) 366 3106
Email: p.pestieau@ulg.ac.be

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=105854

*This Paper has been presented at the CEPR-EPRU workshop ‘Dynamic Aspects of Public Expenditure’. We would like to thank the participants and particularly our discussant, Ken Smetters, for their helpful comments and suggestions. We are also grateful to David de la Croix for his remarks.

Submitted 21 November 2003

1 Introduction

Investment in education by the family is the single most important form of transfer between generations. It is considered as a key source of growth and a crucial factor of inequality. The government can have a number of reasons to intervene in this area. To improve income distribution and foster growth, but also to correct for standard market failures. Intervention can be direct or indirect: taxing or subsidizing investment in education, providing public education or acting on income distribution through income taxation.

Studying the design of educational policy raises a number of modelling issues. First, there is the way parental altruism is specified. There are several possible motivations behind parental involvement in their children's education: an exchange motive (parents expect some reciprocity from their children), pure altruism (parents are concerned by the welfare of their kids), limited altruism (parents educate their children out of some joy of giving). These three motives have different implications in terms of equity, efficiency and growth. Each of them calls for specific public interventions.

Second, when the joy of giving motive is considered as the relevant one, there is the question of how to treat it in the social planner's objective function. Should it be included hence leading to some double counting or should it be laundered out? Third, there is the question of observability. In the tradition of income taxation theory, ability is not observed nor is labor supply. Investment in education by the family may also not be observable at least at the individuals' level.

To deal with these issues, we adopt a model of successive generations. Each generation consists of two types of individuals differentiated according to their ability at work. They work, consume and invest in the education of their children. Their investment contributes to raising the probability that their children will have a high level of ability. Their motivation is the so-called joy of giving or "warm glow" as opposed to pure altruism. By using such an educational technology, we make sure that in each period society consists only of two types (or at least of a finite number). Public policy can affect the relative size of each type and the distribution of disposable income. Individual abilities are not observable. Two assumptions regarding the observability of investment in education are considered. First, we consider the case where human capital investment is observable at the individual level and can thus be subject to a non linear tax. Alternatively, we study the case where educational expenditures are observable only at an aggregate (and anonymous) level and can be subject only to a linear tax. Policy instruments are thus a non linear tax on earnings, public provision of education and a (linear or nonlinear) tax or subsidy on private education.

Our model is inspired by Cremer *et al.* (2002). In that paper the sources of inequality are inherited wealth and productivity which are discrete random variables. The probability that a child receives a high inheritance depends on the parent's investment in a bequest technology. Here we do not have inherited wealth and individuals differ solely in their productivity. For each individual productivity is determined by nature according to a probability distribution which depends on the investment in human capital by his altruistic parents.

Within such a setting, we derive the optimal income tax structure and the formula for the tax/subsidy on private education and for the level of public education. The degree of substitutability between private and public education can be expected to play a crucial role in the design of these public policies.

Anticipating on the results, we show that redistribution mainly rests on the non linear income tax. This is in line with the Atkinson and Stiglitz (1976) proposition. A subsidy or a tax on private investment in human capital of education is generally desirable but the optimal tax rule does not include any term pertaining to redistribution. Instead one has a Pigouvian type term for internalizing the (positive) external effect of education on aggregate welfare and a "merit good" term reflecting the possibility that the warm glow altruism may not be included in the social welfare function (in which case the objective function is not Paretian). These terms are of opposite sign. When the Pigouvian term dominates, a subsidy on education is optimal, otherwise education ought to be taxed. We also provide conditions under which public education is welfare improving and show that the optimal solution may involve total crowding out of private educational expenditures of one type or both types of individuals. Interestingly these properties hold irrespective of the specific information structure considered for education (i.e., both for the case of non linear and the case of linear taxation).

We assume exogenous growth and even more a small open economy. Nevertheless, our problem is related to some contributions in the endogenous growth literature where the growth engine is education, which is either bought by agents (Azariadis and Drazen (1990)) or received from altruistic parents (Glomm and Ravikumar (1992)). There are two common features. First, education has an externality, albeit of a somewhat different nature. In our case this externality concern next generation's aggregate human capital; in the other papers it pertains to the rate of growth of human capital. Second, the specification of altruism is similar. As we do, Glomm and Ravikumar (1992) but also Galor and Zeira (1993) and Aghion and Bolton (1997) adopt a myopic bequest motive rather than a dynastic one.

Observe that Benabou (2002) is an exception in that respect. He studies the effects of progressive income taxation and education finance in a dynamic model and focuses on the trade offs between redistribution, growth and efficiency. His approach like most of the literature on income distribution dynamic is mainly of positive nature. On the normative side there is De Fraja (2002) who studies the education policy chosen by a utilitarian government. He shows that an optimal policy can increase the spread between educational achievements. Like in our paper redistribution is entrusted to optimal income taxation whereas the purpose of educational policy is to increase the size of the pie to be divided as fairly as possible.

2 The model

Consider a model with successive generations. Each individual is characterized by a level of productivity which can only take two values. Individuals draw utility from consumption, from leisure and from their investment in an education technology which affects their children's human capital. More precisely, their investment increases the probability that their children have a high productivity. We assume a small open economy which means that both the interest rate and the wage rate are given and we focus on the steady-state solution.

2.1 The household's problem

All individuals have the same strictly quasi concave utility:

$$U\left(c^i, x^i, L^i\right) = u\left(c^i\right) + h\left(x^i\right) - v\left(L^i\right) \tag{1a}$$

where c^i is consumption, x^i the investment in education and L^i the labor supply. Separability is assumed for two reasons: that of $h(\cdot)$ is to allow for some laundering out of utilities later and that of $v(\cdot)$ is to keep in line with Atkinson and Stiglitz result.

The following instruments of public intervention are considered. First, there is a non linear income tax on $w^i L^i$ where $w^i = w^1$ or w^2 with $w^2 > w^1$. First-best lump-sum taxation is not possible as L^i and w^i are not observable. Second, depending on the information structure, there may be a linear or a non linear tax on x^i, private investment in education. Finally, there may be public education with expenditures $e \geq 0$.

The government does not observe productivity and labor supply; it however observes the relative number of high and low productivity individuals, respectively π^2 and π^1. In other words the distribution of types is known.

The variable π^i is central to our analysis; unlike in traditional optimal tax models, it is endogenous in our setting. At a given period the probability π_t^2 results from investment in human capital. Put differently, the distribution of abilities of generation t depends on the education investment of generation $t-1$. Formally

$$\pi_t^2 = \pi_{t-1}^2 H\left(x_{t-1}^2, e_{t-1}\right) + \pi_{t-1}^1 H\left(x_{t-1}^1, e_{t-1}\right), \tag{1b}$$

where $H\left(x^i, e\right)$ has partial derivatives $H_1^i > 0, H_2^i \geq 0$ and $H_{12}^i \geq 0$. Observe that we can think about $H\left(x, e\right)$ as representing the production technology for human capital, with private and public education as inputs. More precisely, the output produced would be the probability of a child being of high ability. Children of high ability parents of generation $t-1$ then have a probability of $H\left(x_{t-1}^2, e_{t-1}\right)$ of being of high ability. Similarly, children of low ability parents are of high ability with probability $H\left(x_{t-1}^1, e_{t-1}\right)$. The probability that a randomly chosen child from generation t is of high ability is then given by (1b).[1] For notational convenience we shall often use the notation $\pi \equiv \pi^2 = 1 - \pi^1$.

2.2 First-best

As a reference, we start by deriving the first-best optimality conditions. The considered objective is the discounted sum of utilities with a discount factor $\gamma < 1$. To allow for alternative treatments of the altruistic utility term $h\left(x\right)$, we use a parameter ε with $0 \leqslant \varepsilon \leqslant 1$. When $\varepsilon = 0$, the government does not include the joy of giving in its welfare criterion. This is the position advanced by Harsanyi (1995) and Hammond (1987) who have advocated "excluding all external preferences, even benevolent ones, from our social utility function". When $\varepsilon = 1$, the government includes the joy of giving in its objective; this is a pure utilitarist position. Using (1a) and (1b), we

[1] We assume large numbers so that this probability gives us the effective proportion of high ability individuals.

write:

$$\pounds = \sum \gamma^t \left\{ \sum_i^2 \pi_t^i \left[u\left(c_t^i\right) - v\left(L_t^i\right) + \varepsilon h\left(x_t^i\right) \right] \right.$$
$$\left. - \mu_t \sum_i^2 \left[\pi_t^i \left(c_t^i + x_t^i - w^i L_t^i \right) + e_t \right] - \eta_t \left[\pi_{t+1}^2 - \sum \pi_t^i H\left(x_t^i, e_{t-1}\right) \right] \right\}$$

where μ and η are the Lagrangian multipliers associated with the resource constraint and the human capital technology respectively.

Differentiating $\pounds$ with respect to the first-best control variables, x_t^i, c_t^i, L_t^i, π_t^i and $e_t \geq 0$, and evaluating in the steady-state yields the following optimality conditions:

$$x^i \; : \; \varepsilon h'\left(x^i\right) - \mu + \eta H_1\left(x^i, e\right) = 0, \tag{2a}$$
$$c^i \; : \; u'\left(c^i\right) - \mu = 0, \tag{2b}$$
$$L^i \; : \; v'\left(L^i\right) - \mu w^i = 0, \tag{2c}$$
$$\pi \; : \; v^1\left(L^1\right) - v^2\left(L^2\right) - \mu\left[w^1 L^1 - w^2 L^2\right] - \gamma^{-1}\eta = 0, \tag{2d}$$
$$e \; : \; \frac{\partial \pounds}{\partial e} = -\mu + \eta \sum \pi^i H_2\left(x^i, e\right) \leq 0, \tag{2e}$$

where (2e) accounts for the possibility of a binding nonnegativity constraint on e, in which case $e = 0$ and $\partial \pounds / \partial e < 0$. For all other variables an interior solution is assumed. Rearranging, one obtains:

$$x^1 = x^2 \; ; \; c^1 = c^2 \; ; \; v'\left(L^i\right) = u'\left(c^i\right) w^i \; ; \tag{3}$$

$$\eta = \gamma\left[\mu\left(w^2 L^2 - w^1 L^1\right) - \left(v\left(L^2\right) - v\left(L^1\right)\right)\right] ; \tag{4}$$

$$\frac{\partial \pounds}{\partial e} = -\varepsilon \sum \pi^i h'(x^i) + \eta \sum \pi^i H_1(x^i, e) \left[\frac{H_2(x^i, e)}{H_1(x^i, e)} - 1 \right] \leq 0 \tag{5}$$

Conditions (3) are rather standard ones. With a utilitarian objective and separable utility functions, consumption and saving are type-independent and the more able work more than the less able. Condition (4) provides the expression for the multiplier (shadow price) associated with the probability of being productive. Roughly speaking, this measures the contribution to social welfare from turning a less productive individual into a more productive individual. The productive individual has a higher output (w^2L^2 rather than w^1L^1), but also works more, leading to a higher disutility of labor ($v\left(L^2\right)$ as opposed to $v\left(L^1\right)$).[2]

Finally, equation (5) characterizes the optimal level of public education. Consider the case where e and x are perfect substitutes so that $H(x,e) = H(x+e)$ which implies that the second term on the RHS of (5) vanishes. Then there is no reason to have any public education; we necessarily have $\partial\pounds/\partial e \leq 0$ at $e = 0$. When $\varepsilon = 0$, the optimal level of $x+e$ can be achieved with any combination of the two instruments (and we have $\partial\pounds/\partial e = 0$).[3]. However, when $\varepsilon > 0$, x brings some additional social welfare through the *warm glow* effect (and we have $\partial\pounds/\partial e < 0$). Either, way, when x and e are perfect substitutes in H, public education is never welfare improving.

When the two forms of education are not perfect substitutes, $e > 0$ is of course possible, as long as the second term of the RHS of (5) is positive (at $e = 0$) and outweights the first term. This is true particularly when public education is a necessary input in the human capital "production technology", i.e., when $H_2\left(x^i, 0\right) = \infty$. More generally, the second term favors public education when the marginal rate of technical substitution between e and x (i.e., $H_2(x^i,0)/H_1(x^i,0)$) is greater than one (the marginal cost of e).

2.3 *Laissez-faire* and decentralization

In a decentralized economy with an income tax function $T(wL)$ and a consumer price for x^i equal to p, the problem for each individual of type i is to maximize:

[2] These terms are multiplied by the dicount factor because there is a one period lag between input (x or e) and output (π) in the education technology.

[3] As long as e is smaller than the optimal level of $x+e$.

$$u\left[w^iL^i - T\left(w^iL^i\right) - px^i\right] - v\left(L^i\right) + h\left(x^i\right).$$

This yields the FOC:

$$u'\left(c^i\right)w^i\left(1 - T_i'\right) = v'\left(L^i\right) \tag{6}$$

and

$$u'\left(c^i\right)p = h'\left(x^i\right). \tag{7}$$

In a pure *laissez-faire*, $T' = 0$ and $p = 1$; there is no income tax and no tax or subsidy on x. Income and consumption levels differ between types and x^i's are determined according to (7) with $p = 1$. This however does not yield the first-best optimum. Leaving public education aside for the time being, there are two sources of sub-optimality. The first is of distributional nature and due to the use of a utilitarian social welfare function. Under full information, which we assume for the time being, the decentralization of the first-best optimum, then requires lump-sum taxes to achieve conditions (3). The second problem is that x^i's are not determined according to the appropriate tradeoff (specified by (2a)–(2b)). Consequently, the decentralization of the first-best also requires a (positive or negative) tax on x^i's. Denoting the per-unit subsidy or tax on private education by τ it follows from (2a)–(2b) and (7) that the decentralization requires:

$$1 + \tau = p = \frac{h'\left(x^i\right)}{u'\left(c_i\right)} = \left[\frac{h'\left(x^i\right)}{\varepsilon h'\left(x^i\right) + \eta H_1\left(x^i, e\right)}\right] = \left[\varepsilon + \eta\frac{H_1\left(x^i, e\right)}{h'\left(x^i\right)}\right]^{-1}, \tag{8}$$

or equivalently:

$$\tau = -\frac{\eta}{\mu}H_1\left(x^i, e\right) + \frac{1-\varepsilon}{\mu}h'\left(x^i\right), \tag{9}$$

Observe that the tax is linear (the same rate applies to x^1 and to x^2); this is because we are decentralizing a utilitarian optimum at which all consumption levels are equalized.

The first term in brackets gives the ratio between private benefits of educational spending (warm glow) and social benefits (part of warm glow accounted for in welfare plus social value for the future generation). The social value of x^i for future generations is equal to η, the shadow price of π (determined by (4)) times H_1, the induced increase in π. This term reflects the educational externality: individuals only see the "joy of giving" benefit from their investment and not its social value for the future generation. This tends to reduce the tax or may even call for a subsidy. The parameter ε measures the weight the social planner gives to the joy of giving. Thus if $\varepsilon = 1$, $\tau < 0$ holds for sure. However if the social planner does not include the joy of giving in its welfare measure, then a tax is not impossible. Why a tax? Putting aside the externality term for the time being, if $\varepsilon = 0$, the social planner can tax x at no welfare cost. The level of x does not directly appear in the social welfare function and taxing it, it is a good source of non distortionary revenue. More generally, whenever $\varepsilon < 1$ the planner puts less weight on the warm-glow term than an individual. The planner has "paternalistic" (non Paretian) preferences which tend to favor a taxation of x.[4]

3 Second-best taxation

We now introduce second-best taxation, namely a non linear tax on earnings and a linear or non-linear tax on private education. We also introduce public education in (1b) but keep the specification (1a). We need some additional notation: R_t^i denotes disposable income; I_t^i before tax earnings.

[4] In the special case where $H(x,e) = h(x) + h(e)$, (8) reduces to

$$p = [\varepsilon + \eta]^{-1}.$$

There is a subsidy ($\tau < 0$) if $\varepsilon + \eta > 1$ and a tax otherwise.

We are now in a setting wherein the government does not observe the w^i; thus in designing its tax policy, it has to make sure that high productivity individuals do not mimic low productivity individuals to pays less tax. In the case of non linear taxation of x^i, the government observes both c^i and x^i. Then it controls R^i, I^i, e, c^i and x^i. In the case of linear taxation of x^i, it controls R^i, I^i, e and the price of x^i, p. Then we use the supply function x^i with p and R^i as arguments (from (7)).

3.1 Non linear taxation of private education

The Lagrangian of this problem can be written as:

$$\begin{aligned}
\pounds_1 &= \sum_t^{\infty} \gamma^t \left\{ \sum_i^2 \pi_t^i \left[u\left(c_t^i\right) - v\left(\frac{I_t^i}{w^i}\right) + \varepsilon h\left(x_t^i\right) \right] \right. \\
&\quad -\mu_t \left(\sum_i^2 \pi_t^i \left[c_t^i + x_t^i - I_t^i\right] - e_t \right) \\
&\quad +\lambda_t \left(u\left(c_t^2\right) - v\left(\frac{I_t^2}{w^2}\right) + h\left(x_t^2\right) - u\left(c_t^1\right) + v\left(\frac{I_t^1}{w^2}\right) - h\left(x_t^1\right) \right) \\
&\quad \left. -\eta_t \left(\pi_{t+1}^2 - \sum \pi_t^i H\left(x_t^i, e_t\right) \right) \right\}.
\end{aligned}$$

where μ, λ and η are the multiplier associated with the revenue constraint, the self-selection constraint and the definition of π^2 respectively. We derive the following FOC:

$$\frac{\partial \pounds_1}{\partial c_t^1} = \gamma^t \left\{ \pi_t^1 u'\left(c_t^1\right) - \mu_t \pi_t^1 - \lambda_t u'\left(c_t'\right) \right\} = 0 \tag{10a}$$

$$\frac{\partial \pounds_1}{\partial c_t^2} = \gamma^t \left\{ \pi_t^2 u'\left(c_t^2\right) - \mu_t \pi_t^2 + \lambda_t u'\left(c_t^2\right) \right\} = 0 \tag{10b}$$

$$\frac{\partial \pounds_1}{\partial x_t^1} = \gamma^t \left\{ \pi_t^1 \varepsilon h'\left(x_t^1\right) - \mu_t \pi_t^1 - \lambda_t h'\left(x_t^1\right) + \eta_t \pi_t^1 H_1\left(x_t^1, e_t\right) \right\} = 0 \tag{10c}$$

$$\frac{\partial \pounds_1}{\partial x_t^2} = \gamma^t \left\{ \pi_t^2 \varepsilon h'\left(x_t^2\right) - \mu_t \pi_t^2 + \lambda_t h'\left(x_t^2\right) + \eta_t \pi_t^2 H_1\left(x_t^2, e_t\right) \right\} = 0 \tag{10d}$$

$$\frac{\partial \pounds_1}{\partial I_t^1} = -\gamma^t \left\{ \pi_t^1 v'\left(\frac{I_t^1}{w^1}\right) \frac{1}{w^1} - \mu_t \pi_t^1 - \lambda v'\left(\frac{I_t^1}{w^2}\right) \frac{1}{w^2} \right\} = 0 \tag{10e}$$

$$\frac{\partial \pounds_1}{\partial I_t^2} = -\gamma^t \left\{ \pi_t^2 v'\left(\frac{I_t^2}{w^2}\right) \frac{1}{w^2} - \mu_t \pi_t^2 - \lambda v'\left(\frac{I_t^2}{w^2}\right) \frac{1}{w^2} \right\} = 0 \tag{10f}$$

$$\frac{\partial \pounds_1}{\partial e_t} = -\gamma^t \left\{ \mu_t - \eta_t \sum \pi_t^i H_2\left(x_t^i, e_t\right) \right\} \leq 0. \tag{10g}$$

By rearranging theses FOC in the steady-state, we obtain:

$$\frac{v'\left(L^2\right)}{u'\left(c^2\right)w_2} = 1; \tag{11a}$$

$$\frac{v'\left(L^1\right)}{u'\left(c^1\right)w^1} = 1 + \frac{\lambda}{\mu\pi^1}\left[v'\left(\frac{I^1}{w^2}\right)\frac{1}{w^2} - v'\left(\frac{I^1}{w^1}\right)\frac{1}{w^1}\right] \tag{11b}$$

$$\frac{h'\left(x^i\right)}{u'\left(c^i\right)} = 1 - \frac{\eta}{\mu}H_1\left(x^i, e\right) + \frac{1-\varepsilon}{\mu}h'\left(x^i\right); \qquad (i = 1, 2) \tag{11c}$$

The first two FOC are standard conditions of optimal income taxation with two types: no distortion at the top for type 2; positive marginal tax for type 1. We shall now successively study the optimal tax (or subsidy) on private education and the appropriate amount of public education (if any).

3.1.1 The optimal tax or subsidy on private education

Using (7) and defining the τ^i as individual i's marginal (positive or negative) tax on education we can write (11c) as

$$\tau^i = -\frac{\eta}{\mu}H_1\left(x^i, e\right) + \frac{1-\varepsilon}{\mu}h'\left(x^i\right), \tag{12}$$

This expression is exactly equivalent to (9), which gives the first-best trade-off. This finding is consistent with Atkinson and Stiglitz proposition and results from the separability properties of our utility function. Because of this separability, the incentive constraint cannot be relaxed by distorting the choices of some individuals. However, unlike in a conventional Atkinson and Stiglitz setting, we do not obtain zero, nor even uniform taxes on x here. The tax rate defined by (12) will in general differ between individuals. This is because x^i's are not equalized at the second best solution (while they were at the utilitarian optimum).

Keeping in mind that the tax rate is now in general type specific, we obtain otherwise the same optimal tax *rules* as in the first-best. In particular, without externality ($\eta = 0$) and without laundering out ($\varepsilon = 1$), there is no distortion in the choice of private education. The presence of externality implies subsidizing private education; laundering out individual utilities

($\varepsilon < 1$) on the other hand implies taxing education. In other words, one has a Pigouvian type term for internalizing the (positive) external effect of education on aggregate welfare and a "merit good" term reflecting the possibility that the warm glow altruism may not be included in the social welfare function (in which case the objective function is not Paretian). These terms are of opposite sign. When the Pigouvian terms dominates, a subsidy on education is optimal, otherwise education ought to be taxed. Those results are pretty intuitive and consistent with those obtained in the first-best.

3.1.2 The second-best level of public education

Let us now examine whether there is a role for public education in the optimal policy mix. Using (10c) and (10d) to simplify the steady-state version of (10g) yields:

$$\frac{\partial \pounds_1}{\partial e} = \gamma \left\{ \lambda[h'(x^1) - h'(x^2)] - \varepsilon[\pi^2 h'(x^2) + \pi^1 h'(x^1)] + \eta \overline{H}_1 \left[\frac{\overline{H}_2}{\overline{H}_1} - 1 \right] \right\}, \tag{13}$$

where $\overline{H}_k = \sum_i \pi^i H_k(x^i, e)$ is the average level of H_k ($k = 1, 2$).

Let us once again start with the case where x and e are perfect substitutes, that is when $H(x^i, e) = H(x^i + e)$. Recall that there is then no need for public education in a first-best setting; see Section 2.2. In a second-best setting, public education has a more complex impact. With perfect substitutes, the third term on the RHS of (13) vanishes. The first term on the RHS of (13) is positive by the concavity of h. The second term (accounting for the negative sign) is negative. Observe that since we have used (10c) and (10d) to derive (13), this expression is valid only as long as we have an interior solutions for x^1 and x^2.

Now consider the case where $\varepsilon = 0$ (complete laundering out). Under this assumption, the second term on the RHS vanishes and we have $\partial \pounds_1 / \partial e > 0$ as long as we have an interior solution for private education spending x^1 and x^2. Consequently, for $\varepsilon = 0$ it is always optimal to have a positive level of public education. Furthermore, the level of e has to be increased until at least one of the private education terms is totally crowded out (i.e., we when e is sufficiently large to yield $x^1 = 0$). This result appears surprising at first because with perfect substitutes, an increase in e has exactly the same impact on probabilities and on the budget constraints as a uniform increase of x^1 and x^2. However, the crucial difference is that e does not appear in the incentive constraints. A uniform increase in x^1 and x^2 would violate the incentive constraint while a uniform decrease would relax it. Consequently, an increase in e along with a uniform decrease in x^1 and x^2 is welfare improving. It relaxes the incentive constraint by $h'(x^1) - h'(x^2)$, which when multiplied by λ, the shadow price of the incentive constraints yields the first term in the RHS of (13).

We thus have a rather surprising result. While there is no redistributive role for taxing or subsidizing private education (because of the Atkinson and Stiglitz property) here, public education may be an effective instrument in redistributive policy. This is because it relaxes an otherwise binding incentive constraint.

When $\varepsilon > 0$, the effect just discussed continues to be at work. However, substituting e for x now has a direct welfare cost because e does not produce a warm glow effect and because the warm glow term now contributes towards social welfare. To make this more explicit we can use (12) while defining $\overline{\tau} = \sum_i \pi^i \tau^i$ (the average of the individuals marginal tax rates) to rewrite (10g) (again in the case of perfect substitutes):

$$\frac{\partial \pounds_1}{\partial e} = -\gamma \sum_i \pi^i \left[h'\left(x^i\right) \frac{1-\varepsilon}{\overline{\tau}} - \eta \frac{\overline{\tau}+1}{\overline{\tau}} H'\left(x^i + e\right) \right]. \tag{14}$$

With $\varepsilon = 1$ we have from (12) $\tau^i < 0$ for both types so that $\overline{\tau} < 0$ which from (14) yields $\partial \pounds_1 / \partial e < 0$ and thus $e = 0$. Consequently, when there is no laundering out and when the two types of education are perfect substitutes public education is not desirable.

Finally, imperfect substitution between x and e has a similar impact as in a first-best setting. We can think of $\overline{H}_2 / \overline{H}_1$ as the aggregate marginal technical rate of substitution between e and x (a uniform increase of x^1 and x^2). When this rate of substitution exceeds one, efficient production of H calls for a positive level of public eduction. The third term on the RHS of (13) is then positive making a positive e more likely.[5]

3.2 Linear taxation of private education

We now turn to the case of a linear subsidy or tax on private education. This corresponds to a setting where x^i's are not observable at the individual level. However, aggregate (and anonymous) transactions are observable and can be subject to a linear tax at rate τ_t yielding a price p_t which is the same for all types. The income tax, on the other hand, continues to be non linear (before tax income is observable at an individual level). The Lagrangian of the government's problem is then given by:

$$
\begin{aligned}
\pounds_2 \quad = \quad & \sum_t \gamma^t \left\{ \sum_i^2 \pi_t^i \left[u\left(R_t^i - p_t x\left(p_t, R_t^i\right)\right) - v\left(\frac{I_t^i}{w^i}\right) \right.\right. \\
& \left. +\varepsilon\, h\left(x\left(p_t, R_t^i\right)\right)\right] + \mu_t \sum_i^2 \pi_t^i \left[I_t^i - R_t^i + (p_t - 1)\, x\left(p_t, R_t^i\right) - e_t\right] \\
& +\lambda_t \left[u\left(R_t^2 - p_t x\left(p_t, R_t^2\right)\right) - v\left(\frac{I_t^2}{w^2}\right) + h\left(x\left(p_t, R_t^2\right)\right) \right. \\
& \left. - \left(u(R_t^1 - p_t x\left(p_t, R_t^1\right)\right) + v\left(\frac{I_t^1}{w^2}\right) - h\left(x\left(p, R_t^1\right)\right)\right)\right] \\
& \left. -\eta_t \left[\pi_{t+1}^2 - \sum_i^2 \pi_t^i H\left(x\left(p_t, R_t^i\right), e_t\right) \right] \right\}.
\end{aligned}
$$

where $x\left(p_t, R_t^i\right)$ is demand for x, with R_t^i denoting after tax (disposable) income while before tax income is I_t^i. Observe that not being observable, educational expenditures can no longer be *directly* controlled. They can only be controlled indirectly through p (and R^i). We obtain the following

[5] We also have $\gamma\eta(\overline{H}_2 - \overline{H}_1)$ as an additional, positive term in (14).

FOC:

$$\frac{\partial \pounds_2}{\partial R_t^1} = \gamma^t \left\{ \pi_t^1 \left[u'\left(c_t^1\right) + (\varepsilon - 1)\, h'\left(x_t^1\right) \frac{\partial x}{\partial R_t^1} \right] - \mu_t\, \pi_t^1 \left[1 - \tau_t \frac{\partial x}{\partial R_t^1} \right] \right.$$

$$\left. -\lambda\, u'\left(c_t^1\right) + \eta_t\, \pi_t^1 H_1\left(x_t^1, e_t\right) \frac{\partial x}{\partial R_t^1} \right\} = 0, \tag{15a}$$

$$\frac{\partial \pounds_2}{\partial R_t^2} = \gamma^t \left\{ \pi_t^2 \left[u'\left(c_t^2\right) + (\varepsilon - 1)\, h'\left(x_t^2\right) \frac{\partial x}{\partial R_t^2} \right] - \mu_t\, \pi_t^2 \left[1 - \tau_t \frac{\partial x}{\partial R_t^2} \right] \right.$$

$$\left. +\lambda_t\, u'\left(c_t^2\right) + \eta_t\, \pi_t^2 H_1\left(x_t^2, e_t\right) \frac{\partial x}{\partial R_t^2} \right\} = 0, \tag{15b}$$

$$\frac{\partial \pounds_2}{\partial p_t} = \gamma^t \left\{ \sum \pi_t^i \left[-u'\left(c_t^i\right) x_t^i + (\varepsilon - 1)\, h'\left(x_t^i\right) \frac{\partial x\left(p_t, R_t^i\right)}{\partial p_t} \right\} \right.$$

$$+\mu_t \sum \pi_t^i \left[x_t^i + (p_t - 1) \frac{\partial x\left(p_t, R_t^i\right)}{\partial p_t} \right] - \lambda_t \left[u'\left(c_t^2\right) x_t^2 - u'\left(c_t^1\right) x_t^1 \right]$$

$$\left. +\eta_t \sum \pi_t^i H_1\left(x_t^i, e_t\right) \frac{\partial x\left(p_t, R_t^i\right)}{\partial p_t} \right] \right\} = 0. \tag{15c}$$

The expression for the first-order condition with respect to e_t continues to be given by (10g). Rearranging these three FOC and taking the steady-state values, one obtains:

$$p - 1 = \frac{\sum \pi^i \frac{\partial\, \tilde{x}\left(p, R^i\right)}{\partial p} \left[(1 - \varepsilon)\, h'\left(x^i\right) - \eta H_1\left[x^i, e\right] \right]}{\mu \sum \pi^i \frac{\partial\, \tilde{x}\left(p, R^i\right)}{\partial p}} \tag{16}$$

where the tilde denotes compensated derivatives.

Equation (16) is to be compared with equation (11c). The difference is that now there is a single instrument (namely p) controlling both x^1 and x^2. In (11c) $h'\left(x^i\right)/u'\left(c^i\right)$ corresponds to an individualized price on x^i. Equation (16) has the same components but averaged over the two types of individuals. In the numerator the externality term pushes for a subsidy and the possibility of laundering out pushes for a tax on x. The denominator reflects the efficiency cost of linear taxation; it is relatively low when the compensated derivative is small. Denoting

$$\varphi^i = \pi^i \frac{\partial\, \tilde{x}\left(p, R^i\right)}{\partial p} \Bigg/ \sum \pi^i \frac{\partial\, \tilde{x}\left(p, R^i\right)}{\partial p}\,, \tag{17}$$

we have:

$$p-1=\sum\varphi^{i}h'\left(x^{i}\right)\frac{1-\varepsilon-\eta\frac{H_{1}\left(x^{i},e\right)}{h'\left(x^{i}\right)}}{\mu}. \tag{18}$$

For $\varepsilon = 1$ there is an unambiguous case for subsidizing investment in education; for $\varepsilon = 0$ this is not clear anymore.[6]

To study the desirability of public education, we take (10g) in the steady-state along with (16). For simplicity we restrict ourselves to the case of perfect substitutes: $H\left(x^{i},e\right)=H\left(x^{i}+e\right)$ with first derivative $H'\left(x^{i}+e\right)$.[7] Then using (17):

$$\frac{\partial \pounds_{2}}{\partial e}=-\gamma\sum_{i}\left[\varphi^{i}h'\left(x^{i}\right)\frac{1-\varepsilon}{p-1}-\eta\left(\frac{\varphi^{i}}{p-1}+\pi^{i}\right)H'\left(x^{i}+e\right)\right]. \tag{19}$$

If we further assume that the compensated derivatives are equal for both types of households[8] we have $\varphi^{i}=\pi^{i}$ and (19) can be rewritten as:

$$\frac{\partial \pounds_{2}}{\partial e}=-\gamma\sum_{i}\pi^{i}\left[h'\left(x^{i}\right)\frac{1-\varepsilon}{p-1}-\eta\frac{p}{p-1}H'\left(x^{i}+e\right)\right], \tag{20}$$

keeping in mind that $p-1<0$ for ε sufficiently large. Equation (20) is the counterpart for the linear case to (14) and it can be interpreted accordingly. When there is no laundering out ($\varepsilon = 1$), the government has no reason to push for public education since private education generates a "double dividend" (the warm glow effect). Consequently, in case of perfect substitutability we have $e = 0$. When there is a full laundering out, on the other hand, there is a good case for taxing private education ($p-1>0$) and then

[6] If $H\left(x^{i},e\right)=h\left(x^{i}\right)+h\left(e\right)$ which is admittedly a strong assumption, the formula for $p-1$ could be further simplified:

$$p-1=\sum\varphi^{i}h'\left(x^{i}\right)\left[\frac{1-\varepsilon-\eta}{\mu}\right].$$

[7] When e and x are not perfect substitutes, there is an additional term in the expressions, exactly like in the non-linear tax case.

[8] This essentially amounts to assuming that compensated demand function are linear.

some public education can be desirable in case of perfect substitutability.[9] Finally, when the two types of education are not perfect substitutes, the case for public education becomes naturally stronger.

4 Conclusion

We can now draw the two main lessons of this paper. First, Atkinson-Stiglitz proposition holds here. Indirect taxation (or subsidy) is desirable not for redistributive reasons but for correcting two external effects working in opposite directions. In that respect we are close to paper where corrective indirect taxes are used in case of externalities (Cremer *et al.*, 1998). Second we show that public education may be desirable and even crowd out private education when both types of education are perfect substitutes and public education contributes to relax the self-selection constraint. This latter result is only verified when private education is subsidized non-linearly. With linear subsidy one has also some sufficient conditions for positive public education, but not as straightforward as in the non-linear case.

[9] We do not get an expression like (13) for the linear case. This is because the policy variation it reflects (namely and increase in e along with a uniform decrease in x^1 and x^2) is not feasible in the linear context because x is controlled only indirectly.

References

[1] Aghion Ph. and P. Bolton (1997), A trickle down theory of growth and development with debt overhang, *Review of Economics Studies*, 64, 151-172.

[2] Atkinson, A. and J. Stiglitz (1976), The design of the tax structure: direct versus indirect taxation, *Journal of Public Economics*, 6, 55–75.

[3] Azariadis, C. and A. Drazan, (1990), Threshold externalities in economic development, *Quarterly Journal of Economics,* 105, 501–526.

[4] Benabou R. (2002), Tax and education policy in a heterogenous-agent economy: what levels of redistribution maximize growth and efficiency?, *Econometrica*, 70, 481-517.

[5] Cremer, H., P. Pestieau and J.-Ch. Rochet, (2002), Capital income taxation when inherited wealth is not observable, *Journal of Public Economics*, forthcoming.

[6] Cremer, H, F. Gahvari and N. Ladoux, (1998), Externalities and optimal taxation, *Journal of Public Economics*, 70, 343–364.

[7] De Fraja G. (2002), The design of optimal educational policies, *Review of Economic Studies,* forthcoming.

[8] Fernandez, R. and R. Rogerson, (1996), Income distribution, communities and the quality of public education, *Quarterly Journal of Economics*, 111, 135-164.

[9] Galor O. and J. Zeira (1993), Income distribution and macroeconomics, *Review of Economic Studies*, 60, 35-52

[10] Glomm, G. and B. Ravikumar, (1992), Public versus private investment in human capital: endogenous growth and income inequality, *Journal of Political Economy*, 100, 818-834.

[11] Hammond, P., [1987], Altruism, in: *The New Palgrave: A Dictionary of Economics,* Macmillan, London.