Intergenerational transfer of human capital and optimal education policy[1]

Helmuth Cremer
University of Toulouse (IDEI and GREMAQ),

Pierre Pestieau

CREPP, University of Liège, CORE and Delta

This version February 2005

Abstract

We study the design of education policies (subsidies and public education) when parents' investment in education is motivated by warm glow altruism and determines the probability that a child has a high ability. The optimal subsidy is not necessarily positive. It is determined by two conflicting terms: a Pigouvian term (warm glow altruists do not properly account for the impact of education on future generations) and a "paternalistic" effect (the warm glow term may not be fully included in social welfare). Finally, total crowding out of private expenditure (for one of the types) by public education may be desirable.

Running head: Intergenerational transfer of human capital

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 for 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, on the other hand, may also not be observable.

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

1

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 or subsidy. 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 or subsidy. 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.* (2003). 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 that 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 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 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 social 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

2

of the specific information structure considered for education (i.e., both under non-linear and under linear taxation).

Throughout the paper, our main focus lies on normative issues. From that perspective this paper differs significantly from most of the existing literature on education finance and income distribution in dynamic settings which is of positive nature. Furthermore, many existing contributions fall in the endogenous growth literature and consider education as the growth engine. We assume exogenous growth (a small open economy) and we concentrate on the steady state. In other words, the growth feature of our model is negligible. Nevertheless, it is worth citing some endogenous growth contributions. In that literature education is either bought by agents (Azariadis and Drazen (1990)) or received from altruistic parents (Glomm and Ravikumar (1992)). There are two common features with our model. First, education has an externality, albeit of a somewhat different nature. In our case this externality concerns 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.[1] An exception is Benabou (2002) who considers pure altruism. He uses a setting with infinitely lived individuals (dynasties) to study the effects of progressive income taxation and education finance in a dynamic model and focuses on the trade offs between redistribution, growth and efficiency.

On the normative side De Fraja (2002) studies the education policy chosen by a utilitarian government. He shows that an optimal policy can be "elitist" in the sense that it increases the spread between educational achievements. As in our paper, government policies include an optimal income tax along with education subsidies. Nevertheless, the setting differs in many respect from ours in particular with regard to the education technology which is a main feature of our setting. In particular, De Fraja (2002) does not distinguish between private and public education (as separate inputs). Furthermore, in our setting the way education in a given period affects the skill distribution in the

---

[1]Another positive study is provided by Glomm and Kaganovich (2003). They consider a setting with uniform and universal public expenditures on education in which the government also runs a pay-as-you-go social security system. They study the impact of these policies on inequality.

next period leads to an optimal income tax problem with endogenous proportions of types; De Fraja (2002) has no such feature in his setting.

The normative focus of our paper also explains why we take a view which may appear to be at odds with financing arrangements of education in the real world. In many countries education is to a large extend publicly provided whereas our model focuses, at least initially, on privately provided education. Though empirically of little relevance, the private education setting is an important benchmark against which the role for public education has to be assessed.

2 The model

Consider a model with successive generations. Each individual is characterized by a level of productivity that can only take two values. Individuals draw utility from consumption, from leisure and from their investment in an education technology that affects their children's human capital. More precisely, their investment increases the probability that their children have 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 $h\left(\cdot\right)$ is to allow for some laundering out of utilities later and that of $v\left(\cdot\right)$ is to keep in line with the 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 $w^i \in \{w^1, w^2\}$, 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 expenditures $e \geq 0$.

4

While the government does not observe productivity, it does observe 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 in the previous period. 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 $H\left(x, e\right)$ has partial derivatives $H_1 > 0, H_2 \geq 0$ and $H_{12} \geq 0$. 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).[2] 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 adjusted 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".[3] When $\varepsilon = 1$, the government includes the joy of

---

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

[3] The quote is from Hammond (1987) p. 87.

giving in its objective; this is a pure utilitarian position. Using (1a) and (1b), we write:

$$\mathcal{L} = \sum_{t=0}^{\infty} \gamma^t \left\{ \sum_{i=1}^{2} \pi_t^i \left[u\left(c_t^i\right) - v\left(L_t^i\right) + \varepsilon h\left(x_t^i\right) \right] \right.$$
$$\left. + \mu_t \sum_{i=1}^{2} \left[\pi_t^i \left(w^i L_t^i - x_t^i - c_t^i\right) - e_t \right] - \eta_t \left[\pi_{t+1}^2 - \sum_{i=1}^{2} \pi_t^i H\left(x_t^i, e_t\right) \right] \right\}$$

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

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

$$x^i \quad : \quad \varepsilon h'\left(x^i\right) - \mu + \eta H_1\left(x^i, e\right) = 0, \tag{2a}$$

$$c^i \quad : \quad u'\left(c^i\right) - \mu = 0, \tag{2b}$$

$$L^i \quad : \quad v'\left(L^i\right) - \mu w^i = 0, \tag{2c}$$

$$\pi \quad : \quad v\left(L^1\right) - v\left(L^2\right) - \mu\left[w^1 L^1 - w^2 L^2\right] - \gamma^{-1}\eta = 0, \tag{2d}$$

$$e \quad : \quad \frac{\partial \mathcal{L}}{\partial e} = -\mu + \eta \sum_{i} \pi^i H_2\left(x^i, e\right) \leq 0, \tag{2e}$$

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

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

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

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

Conditions (3) are rather standard. 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 one. The productive individual has a higher output ($w^2 L^2$ rather than

$w^1 L^1$), but also works more, leading to a higher disutility of labor ($v\left(L^2\right)$ as opposed to $v\left(L^1\right)$).[4]

Finally, equation (5) characterizes the optimal level of public education. An interesting case obtains when e and x are perfect substitutes so that $H\left(x,e\right) = H\left(x+e\right)$ 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 $\partial\pounds/\partial e = 0$).[5] 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 $H_2\left(x^i, 0\right) = \infty$. This would, however, be a strong and debatable assumption. 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 $T\left(wL\right)$ and a consumer price for x^i equal to p, the problem for each individual of type i is to maximize:[6]

---

[4] 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.

[5] This is so as long as e is smaller than the optimal level of $x+e$.

[6] The first-best allocation satisfies the resource constraint

$$\sum_{i=1}^{2} \left[\pi_t^i\left(w^i L_t^i - x_t^i - c_t^i\right) - e_t\right].$$

Summing up the individual's budget constraints $c^i = w^i L^i - T\left(w^i L^i\right) - px^i$ and combining with the resource constraints then yields

$$\sum_{i=1}^{2} \pi_t^i[T(w^i L_t^i) + (p-1)x_t^i] = e_t,$$

so that total tax revenue exactly covers eduction expenditures (the government budget constraint is satisfied).

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

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 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 in (8) 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 is a good source of non distortionary revenue. in this case, $\tau < 0$, if $-\eta H_1\left(x^i, e\right) + h'\left(x^i\right) < 0$ by (9); i.e., if the warm-glow motive would induce higher educational expenditure than is socially optimal. 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.[7]

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. The level of public education also continues to be a policy instrument. We need some additional notation: R_t^i denotes disposable income and I_t^i before tax earnings.

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 (subject to incentive compatibility). 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)).

---

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

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

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

3.1 Non-linear taxation of private education

The Lagrangian of this problem can be written as:

$$
\begin{aligned}
\mathcal{L}_1 &= \sum_{t=0}^{\infty} \gamma^t \left\{ \sum_{i=1}^{2} \pi_t^i \left[u\left(c_t^i\right) - v\left(\frac{I_t^i}{w^i}\right) + \varepsilon h\left(x_t^i\right) \right] \right. \\
&\quad + \mu_t \left(\sum_{i=1}^{2} \pi_t^i \left[I_t^i - c_t^i - x_t^i \right] - e_t \right) \\
&\quad + \lambda_t \left(u\left(c_t^2\right) - v\left(\frac{I_t^2}{w^2}\right) + h\left(x_t^2\right) - u\left(c_t^1\right) + v\left(\frac{I_t^1}{w^2}\right) - h\left(x_t^1\right) \right) \\
&\quad \left. - \eta_t \left(\pi_{t+1}^2 - \sum_{i=1}^{2} \pi_t^i H\left(x_t^i, e_t\right) \right) \right\}.
\end{aligned}
$$

where μ, λ and η are the multiplier associated with the revenue constraint, the incentive (or self-selection) constraint and the definition of π^2 respectively. The FOC are provided in Appendix A1. Rearranging theses FOC in the steady-state, we obtain:[8]

$$
\frac{v'\left(L^2\right)}{u'\left(c^2\right)w_2} = 1; \tag{10a}
$$

$$
\frac{v'\left(L^1\right)}{u'\left(c^1\right)w^1} = 1 + \frac{\lambda}{\mu\pi^1}\left[v'\left(\frac{I^1}{w^2}\right)\frac{1}{w^2} - v'\left(\frac{I^1}{w^1}\right)\frac{1}{w^1} \right] \tag{10b}
$$

$$
\frac{h'\left(x^i\right)}{u'\left(c^i\right)} = 1 - \frac{\eta}{\mu}H_1\left(x^i, e\right) + \frac{1-\varepsilon}{\mu}h'\left(x^i\right); \qquad (i = 1, 2) \tag{10c}
$$

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 (10c) as

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

---

[8]To obtain (10b) rewrite (A1e)

$$
\pi_t^1 v'\left(\frac{I_t^1}{w^1}\right)\frac{1}{w^1} = \mu_t\pi_t^1 + \lambda v'\left(\frac{I_t^1}{w^2}\right)\frac{1}{w^2},
$$

and subtract $\lambda v'(I^1/w^1)/w^1$ from both sides, and divide by $\mu_t\pi_t^1 = (\pi_t^1 - \lambda_t)u'\left(c_t^1\right)$ (from (A1a)).

This expression is exactly equivalent to (9), which gives the first-best tradeoff. 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 (11) 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 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 (A1c) and (A1d) to simplify the steady-state version of (A1g) yields:

$$\frac{\partial \mathcal{L}_1}{\partial e} = \gamma \left\{ \lambda[h'(x^1) - h'(x^2)] - \varepsilon[\pi^2 h'(x^2) + \pi^1 h'(x^1)] + \eta \overline{H}_1 \left[\frac{\overline{H}_2}{\overline{H}_1} - 1 \right] \right\}, \qquad (12)$$

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 subsection 2.2. In a second-best setting, public education has

11

a more complex impact. With perfect substitutes, the third term on the RHS of (12) vanishes. The first term on the RHS of (12) is positive by the concavity of h. The second term (accounting for the negative sign) is negative. Observe that since we have used (A1c) and (A1d) to derive (12), 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 $\partial \mathcal{L}_1 / \partial e > 0$ 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 (that is, $x^1 = 0$). 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 (12).

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 while public and private education are otherwise perfect substitutes, they are not perceived in the same way by altruistic parents. More precisely, only private education gives rise to the warm glow effect. A shown above, this special feature explains that in our setting public provision can relax an otherwise binding incentive constraint even though the Atkinson property holds. This is ad odds with Cremer and Gahvari's (1997) result that under fairly general conditions (but in the *absence of altruism*), public provision (and in-kind transfers) are redundant in such settings.

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 (11) while defining $\overline{\tau} = \sum_i \pi^i \tau^i$ (the average of the individuals marginal tax rates) to rewrite (A1g) (again in the case of perfect substitutes):

$$\frac{\partial \mathcal{L}_1}{\partial e} = -\gamma \sum_i \pi^i \left[h'\left(x^i\right) \frac{1-\varepsilon}{\overline{\tau}} - \eta \frac{\overline{\tau}+1}{\overline{\tau}} H'\left(x^i+e\right) \right]. \tag{13}$$

With $\varepsilon = 1$ we have from (11) $\tau^i < 0$ for both types so that $\overline{\tau} < 0$ which from (13) yields $\partial \mathcal{L}_1 / \partial e < 0$ and thus $e = 0$.[9] 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 education. The third term on the RHS of (12) is then positive making a positive e more likely.[10]

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. This implies that before tax income I_t^i and after tax (disposable) income R_t^i continues to be observable at an individual level. What is not observable however, is how the individual allocates his disposable income and in particular how much he spends on education. Not being observable, educational expenditures can no longer be *directly* controlled. They can only be controlled indirectly through p and R^i and the government's problem is now stated in terms of individual

---

[9]Rather than deriving this result from (13), one can use (A1a) to show that

$$\lambda = \pi^1 - \frac{\mu \pi^1}{u'(c^1)};$$

and then inspect (12).

[10]We also have $\gamma \eta (\overline{H}_2 - \overline{H}_1)$ as an additional, positive term in (13).

demands functions $x\left(p_t, R_t^i\right)$ (with price and disposable income as arguments). This has a crucial impact on the incentive constraint; when individual 2 mimicks individual 1, he no longer has to match the spending pattern of the mimicked type.

The formal statement of this problem and the FOC are provided in Appendix A2. Rearranging these three FOC and taking the steady-state values, one obtains:

$$p - 1 = \frac{\sum_i \pi^i \dfrac{\partial \tilde{x}\left(p, R^i\right)}{\partial p}\left[(1 - \varepsilon)\, h'\left(x^i\right) - \eta H_1\left[x^i, e\right]\right]}{\mu \sum_i \pi^i \dfrac{\partial \tilde{x}\left(p, R^i\right)}{\partial p}}, \tag{14}$$

where the tilde denotes derivatives of compensated demands.

Equation (14) is to be compared with equation (10c). The difference is that now there is a single instrument (namely p) controlling both x^1 and x^2. In (10c) $h'\left(x^i\right)/u'\left(c^i\right)$ corresponds to an individualized price on x^i. Equation (14) 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

$$\varphi^i = \pi^i \frac{\partial \tilde{x}\left(p, R^i\right)}{\partial p} \left/ \sum_i \pi^i \frac{\partial \tilde{x}\left(p, R^i\right)}{\partial p} \right., \tag{15}$$

we have:

$$p - 1 = \sum_i \varphi^i h'\left(x^i\right) \frac{1 - \varepsilon - \eta \dfrac{H_1\left(x^i, e\right)}{h'(x^i)}}{\mu}. \tag{16}$$

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

To study the desirability of public education, we take (A1g) in the steady-state along with (14). 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 $H'\left(x^i + e\right)$.[11] Then using (15):

$$\frac{\partial \pounds_2}{\partial e} = -\gamma \sum_i \left[\varphi^i h'\left(x^i\right) \frac{1 - \varepsilon}{p - 1} - \eta \left(\frac{\varphi^i}{p - 1} + \pi^i\right) H'\left(x^i + e\right)\right]. \tag{17}$$

---

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

If we further assume that the compensated derivatives are equal for both types of households,[12] we have $\varphi^i = \pi^i$ and (17) can be rewritten as:

$$\frac{\partial \mathcal{L}_2}{\partial e} = -\gamma \sum_i \pi^i \left[h'\left(x^i\right) \frac{1-\varepsilon}{p-1} - \eta \frac{p}{p-1} H'\left(x^i + e\right) \right],\tag{18}$$

keeping in mind that $p - 1 < 0$ for ε sufficiently large. Equation (18) is the counterpart for the linear case to (13) 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 "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 some public education can be desirable in case of perfect substitutability.[13] Finally, when the two types of education are not perfect substitutes, the case for public education naturally becomes stronger.

4 Conclusion

We can now draw the two main lessons of this paper. First, the Atkinson-Stiglitz proposition holds here. Taxation (or subsidization) of *private* education spending is desirable not for redistributive reasons but for correcting two external effects working in opposite directions. In that respect we are close to papers where corrective commodity taxes are used in the presence of externalities (Cremer *et al.*, 1998). Second, we show that *public* education may be desirable in a second best setting even when it is a perfect substitute to private education while public provision of private education is a redundant instrument. This is because while public and private information are otherwise perfect substitutes, they are not perceived in the same way by altruistic parents. More precisely, only private education give rise to the warm glow effect. Interestingly, public education may completely crowd out private education expenditures for some individuals.

---

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

[13]We do not get an expression like (12) for the linear case. This is because the policy variation it reflects (namely an 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.

Throughout the paper, we assume that parents exhibit a form of paternalistic altruism towards their children. It is then not surprising that fiscal policy is operative as shown by Andreoni (1989). Assuming pure altruism à la Barro would lead to a totally different model or rather to different models depending on the informational assumptions. This being said, as long as altruism is restricted to educational investment as opposed to financial bequests, most of the properties of infinitely lived agents models (including the Ricardian equivalence or the zero capital income tax) do not hold anymore (see on this the discussion by Drazen (1978)) and some of our results could very well continue to be valid.

Observe that the "externality" associated with the education investment appears to be closely related to the specification of warm glow altruism. The warm glow term does not account for the "true" effect of education on the next generation and hence the externality problem. However, one has to realize that the externality is not solely due to the warm glow specification. Even with perfect altruism, education has a public good dimension in the sense that it is average (aggregate) rather than individual spending which determines the next generation's productivities. Any given child's welfare then directly depends on his own productivity, but also indirectly on the average productivity through the resource constraint in the optimal tax problem. An increase in a single individual's productivity will then be in part taxed away so that the contribution of productivity to individual and social welfare differs.

Finally let us recall that government policy in this paper is driven by three factors: myopic individual behavior, distributional considerations and the spillover effect of human capital. The first two can be disputed as being subjective. In contrast, the latter can be seen as the most significant and the least questionable one.

Appendix

A1 First-order conditions for $\pounds_1$

Differentiating $\pounds_1$ yields the following FOC

$$\frac{\partial \pounds_1}{\partial c_t^1} = \gamma^t \left\{ \pi_t^1 u'\left(c_t^1\right) - \mu_t \pi_t^1 - \lambda_t u'\left(c_t^1\right) \right\} = 0, \tag{A1a}$$

$$\frac{\partial \pounds_1}{\partial c_t^2} = \gamma^t \left\{ \pi_t^2 u'\left(c_t^2\right) - \mu_t \pi_t^2 + \lambda_t u'\left(c_t^2\right) \right\} = 0, \tag{A1b}$$

$$\frac{\partial \pounds_1}{\partial x_t^1} = \gamma^t \left\{ \pi_t^1 \varepsilon h'\left(x_t^1\right) - \mu_t \pi_t^1 - \lambda_t h'\left(x_t^1\right) + \eta_t \pi_t^1 H_1\left(x_t^1, e_t\right) \right\} = 0, \tag{A1c}$$

$$\frac{\partial \pounds_1}{\partial x_t^2} = \gamma^t \left\{ \pi_t^2 \varepsilon h'\left(x_t^2\right) - \mu_t \pi_t^2 + \lambda_t h'\left(x_t^2\right) + \eta_t \pi_t^2 H_1\left(x_t^2, e_t\right) \right\} = 0, \tag{A1d}$$

$$\frac{\partial \pounds_1}{\partial I_t^1} = -\gamma^t \left\{ \pi_t^1 v'\left(\frac{I_t^1}{w^1}\right) \frac{1}{w^1} - \mu_t \pi_t^1 - \lambda v'\left(\frac{I_t^1}{w^2}\right) \frac{1}{w^2} \right\} = 0, \tag{A1e}$$

$$\frac{\partial \pounds_1}{\partial I_t^2} = -\gamma^t \left\{ \pi_t^2 v'\left(\frac{I_t^2}{w^2}\right) \frac{1}{w^2} - \mu_t \pi_t^2 + \lambda v'\left(\frac{I_t^2}{w^2}\right) \frac{1}{w^2} \right\} = 0, \tag{A1f}$$

$$\frac{\partial \pounds_1}{\partial e_t} = -\gamma^t \left\{ \mu_t - \eta_t \sum \pi_t^i H_2\left(x_t^i, e_t\right) \right\} \leq 0. \tag{A1g}$$

A2 Linear taxation of private education: formal analysis

The Lagrangian of the government's problem is given by:

$$
\begin{aligned}
\pounds_2 = \sum_{t=0}^{\infty} \gamma^t \Bigg\{ & \sum_{i=1}^{2} \pi_t^i \left[u\left(R_t^i - p_t x\left(p_t, R_t^i\right)\right) - v\left(\frac{I_t^i}{w^i}\right) \right. \\
& + \varepsilon\, h\left(x\left(p_t, R_t^i\right)\right) \Big] + \mu_t \sum_{i=1}^{2} \pi_t^i \left[I_t^i - R_t^i + (p_t - 1) x\left(p_t, R_t^i\right) - e_t \right] \\
& + \lambda_t \left[u\left(R_t^2 - p_t x\left(p_t, R_t^2\right)\right) - v\left(\frac{I_t^2}{w^2}\right) + h\left(x\left(p_t, R_t^2\right)\right) \right. \\
& - \left(u(R_t^1 - p_t x\left(p_t, R_t^1\right)) + v\left(\frac{I_t^1}{w^2}\right) - h\left(x\left(p, R_t^1\right)\right) \right) \Big] \\
& - \eta_t \left[\pi_{t+1}^2 - \sum_{i=1}^{2} \pi_t^i H\left(x\left(p_t, R_t^i\right), e_t\right) \right] \Bigg\}.
\end{aligned}
$$

We obtain the following FOC:

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

$$\left. - \lambda\, u'\left(c_t^1\right) + \eta_t\, \pi_t^1 H_1\left(x_t^1, e_t\right) \frac{\partial x}{\partial R_t^1} \right\} = 0, \qquad \text{(A2a)}$$

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

$$\left. + \lambda_t\, u'\left(c_t^2\right) + \eta_t\, \pi_t^2 H_1\left(x_t^2, e_t\right) \frac{\partial x}{\partial R_t^2} \right\} = 0, \qquad \text{(A2b)}$$

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

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

$$\left. + \eta_t \sum_i \pi_t^i H_1\left(x_t^i, e_t\right) \frac{\partial x\left(p_t, R_t^i\right)}{\partial p_t} \right] \right\} = 0. \qquad \text{(A2c)}$$

The expression for the first-order condition with respect to e_t continues to be given by (A1g).

References

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

[2] Andreoni, J. (1989), Giving with impure altruism: Applications to charity and Ricardian equivalence, *Journal of Political Economy,* 97, 1447-58.

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

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

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

[6] Cremer, H., P. Pestieau and J.-Ch. Rochet (2003), Capital income taxation when inherited wealth is not observable, *Journal of Public Economics*, 87, 2475–2490.

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

[8] Cremer, H. and F. Gahvari (1997), In-kind transfers, self-selection and optimal tax policy, *European Economic Review, 41, 97–114.*

[9] De Fraja, G. (2002), The design of optimal educational policies, *Review of Economic Studies,* 69, 437–466.

[10] Drazen, A. (1978), Government debt, human capital and bequests in a life-cycle model, *Journal of Political Economy*, 86, 506-516.

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

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

[13] Glomm, G. and M. Kaganovich (2003), Distributional effects of public education in an economy with public pensions, *International Economic Review,* 44, 917–938.

[14] 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.

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

[16] Harsanyi, J. (1995) A theory of social values and a rule utilitarian theory of morality, *Social Choice and Welfare*, 12, 319–344.