Should Educational Policies be Regressive?*

Daniel Gottlieb,† Humberto Moreira‡

January 13, 2004

Abstract

In this paper, we show that when the government is able to transfer wealth between generations, regressive policies are no longer optimal. The optimal educational policy can be decentralized through appropriate Pigouvian taxes and credit provision, is not regressive, and provides equality of opportunities in education (in the sense of irrelevance of parental income for the amount of education). Moreover, in the presence of default, the optimal policy can be implemented through income-contingent payments.

Keywords: Education, Pigouvian taxes, student loans, redistribution.
JEL Classification: I28, H23, I22, H52.

*We wish to thank Luis Henrique Braido, Carlos E. da Costa, Daniel Ferreira, Andrew Horowitz, Rodrigo Soares, Thierry Verdier and specially Antoine Bommier and Pierre Dubois for helpful comments and suggestions. As usual, all remaining errors are ours.

†Graduate School of Economics, Getulio Vargas Foundation.

‡Graduate School of Economics, Getulio Vargas Foundation. Corresponding Author. Postal Address: Praia de Botafogo, 190, sala 1121, CEP: 22253-900, Rio de Janeiro, RJ, Brazil. E-mail: humberto@fgv.br.

1 Introduction

The role of educational policies in the equalization of opportunities is a widely accepted issue in political debates. However, a remarkable feature of most educational systems in the world is the huge regressivity of spending per students (i.e., children from wealthier families receive more education than those from poorer families).[1] This regressivity of educational systems may indicate either the presence of some trade-off between equity and efficiency or the inefficiency of observed policies.

The existence of a trade-off between redistribution and efficiency in taxation has been known at least since the work of Mirrlees (1971). In the specific case of education, this issue has been previously discussed by Becker (1991) in the context of the parent's decision on the education provided for children with different abilities. Hare and Ulph (1979) find that the optimal educational policies will be egalitarian (in the sense of constant consumption and utility) only for intermediate abilities.

The theoretical literature on optimal educational policies in an asymmetric-information context was pioneered by Ulph (1977) and Hare and Ulph (1979) who extended the optimal taxation approach of Mirrlees (1971) to address the problem of determining the optimal educational and taxation policies jointly when the ability to benefit from education is unobservable. More recently, De Fraja (2002) studied the optimal educational provision in an overlapping-generations model in the presence of externalities and imperfect capital markets. His results suggest that educational policies should be regressive (in the sense that households with brighter children and higher incomes contribute less than those with less bright children and lower incomes) and do not provide equality of opportunities in education (in the sense of the irrelevance of the household's income to the education received by a child). Therefore, the regressivity of educational systems in most countries may actually reflect the optimal educational policies and the provision of equality of opportunities in education may imply a great efficiency loss.

We shall argue that the results obtained by De Fraja (2002) critically rely on a particular restriction on the government's budget constraint: budget is imposed to be balanced with each generation at any time. Since we are considering an overlapping-generations model, the government would usually be able to transfer between generations. Indeed, this is one important feature of public educational systems: older generations contribute to finance the education of younger generations.[2] Thus, it seems reasonable to assume that governments

[1] See, for example, Fernandez and Rogerson (1996), Kozol (1991) or Psacharopoulus (1986). In the United States, this regressivity is reflected in the large disparity of spending per student across districts. Since 43 percent of elementary and secondary education is financed at local level, 49.9 percent is financed at state level, and only 7.1 percent is financed at federal level (2001 Census of Governments), these differences reproduce the inequality of income distribution. Fernandez and Rogerson (1998) and Inman (1978) provided general equilibrium computations of the welfare gains associated with the centralization of educational expenses.

[2] This kind of intergenerational transfers is also available in pay-as-you-go social-security systems where young generations contribute for the benefits of the older generations.

are able to transfer between generations.

However, when transfers between generations are allowed, the optimal educational policy takes a very different form: it achieves first-best welfare (the maximum amount of welfare that could be reached under perfect information) and provides equality of opportunities in education. Moreover, it can be decentralized through appropriate Pigouvian taxes and the provision of credit.

In the decentralized mechanism, first-best welfare is reached through a subsidy on education to correct the externalities, a lump-sum taxation proportional to the average education and the provision of credit (at the market interest rate). Such a mechanism is not regressive (i.e., wealthier households do not contribute less than poorer households and households with brighter children contribute more than those with less bright children) and does not require knowledge of each household's wealth.

Furthermore, when we incorporate the possibility of default, the optimal educational policy can be implemented through income-contingent payments (Krueger and Bowen, 1993; Barr, 1991).

Hence, our results suggest that the inefficiency results obtained at De Fraja (2002) are closely related to the particular government budget constraint. If the government is able to transfer wealth between generations, then the observed inequalities would reflect an inefficiency in educational systems. We also propose an intuitive and informationally less demanding educational policy.

There is also well-established theoretical literature that emphasizes the public-choice perspective of public-education financing rather than focusing on efficiency arguments. In this literature, it is usually assumed that educational provision must be uniform for each neighborhood and the amount of education is decided through majority voting.[3]

The remainder of the paper is organized as follows. Section 2 lays out the framework of the model. The basic structure is the same as De Fraja (2002). In section 3 we present the laissez-faire equilibrium. In section 4, the government-intervention solution is presented. In subsection 4.1, the first-best equilibrium is characterized; then we present the second-best equilibrium (4.2) and the decentralized equilibrium (4.3). In section 5, we characterize the second-best equilibrium when there is possibility of default. Section 6 summarizes the main results of the paper.

2 The Basic Framework

We consider an overlapping generations model where individuals live for 2 periods. In the first one (childhood), the individual receives an education and a bequest. In the second period (adulthood), she works, has a child, consumes, and provides an education and a bequest for her daughter. Each household consists of a parent and a child. There is a continuum of households with measure

[3] This strand of the literature includes De Bartolome (1990), Epple and Romano (1996), Fernandez and Rogerson (1995, 1996), Johnson (1984), Peltzman (1973), and Stiglitz (1974).

normalized to 1 in every period. As we will focus our attention on steady-state equilibria, time subscripts will be omitted.

Each individual's utility function is

$$U = u(c) + x,$$

where c is her consumption and x is the amount of monetary resources available to the child.[4]

Assumption. $u \in C^2$ satisfies

$$u'(c) > 0, \; u''(c) < 0, \; \lim_{c \to 0} u'(c) = +\infty, \; u(0) = 0, \text{ and} \tag{H1}$$

$$u'(c^*) = 1 \text{ for some } c^* \in \Re_{++}.$$

Notice that the quasi-linearity of the utility function implies that c^* is the amount of consumption whose marginal utility is equal to the marginal utility of the child's monetary resources.

There are two ways of transferring wealth to the child: bequest t and higher future wages (through education e). We normalize the interest rate paid on bequests to 1. Education is transformed in future wages through the household-production technology $y(\theta, e; E)$, where $\theta \in [\theta_0, \theta_1]$ is each child's productivity parameter, e is the amount of education and E is the general level of education.[5]

Substituting the two possible ways of transferring wealth to the child, we get

$$x = y(\theta, e; E) + t. \tag{1}$$

The mother's wealth, denoted by Y, is itself a function the predetermined amount of education she received in her childhood. Let $\Gamma \subset \mathbb{R}_+$ be the space of possible wealth levels. We define $h(Y, \tilde{e})$ as the probability function of Y given the educational profile of the previous generation $\tilde{e}$. As the parent's education is predetermined, we omit the term $\tilde{e}$ for notational convenience.

Let k be the monetary cost of a unit of education. We assume that public and private schools provide education at the same cost implying that the actual provider of education is immaterial. Hence, we abstract from the discussion on whether education should be privately or publicly provided (see Lott (1987)).

Then, the household's budget constraint is

$$Y = c + ke + t. \tag{2}$$

[4] The dependence of the parent´s utility function on the child´s wealth rather than on her utility is an usual assumption and greatly simplifies the analysis. Because of its linearity, it implies that the mother is risk-neutral in the wealth left to her daughter.

[5] Notice that in the model presented, it is immaterial if education serves only as a screening device or whether it enhances productivity.

Assumption. $y \in C^2$ satisfies

$$y_e(\theta, e; E) > 0 \text{ (H2)}, \; y_\theta(\theta, e; E) > 0 \text{ (H3)}, \; y_{e\theta}(\theta, e; E) > 0 \text{ (H4)},$$
$$y_E(\theta, e; E) > 0 \text{ (H5)}, \; y(\theta, 0; E) > 0 \text{ (H6)}, \; y_{ee}(\theta, e; E) < 0 \text{ (H7)},$$
$$\lim_{e \to 0} y_e(\theta, e; E) = +\infty \text{ (H8)}, \lim_{e \to \infty} y_e(\theta, e; E) = 0 \text{ (H9)},$$
$$\lim_{e \to \infty} y_E(\theta, e; E) < k \text{ (H10)}.$$

Assumption $(H3)$ states that ability increases the return of education. $(H4)$ means that education increases earnings more for abler individuals (single-crossing condition). Assumption $(H5)$ means that education is a source of positive externalities.[6] This is the most interesting case in the model presented, although the results trivially hold when there is no externality in education (i.e. $y_E(\theta, e; E) = 0$). Assumptions $(H2)$ and $(H7)$ mean that education increases earnings in a decreasing fashion, while $(H6)$ means that someone with no education is still able to earn some positive salary. Assumption $(H10)$ means that the externality is not big enough that, when the amount of education is infinite, the externalities caused by education exceed the cost of education, while the other assumptions are the usual Inada conditions which are helpful for the existence of the equilibria presented in the following sections. $(H8)$ and $(H9)$ are the usual Inada conditions that ensure the interiority of the solution.

Let $\phi(\theta)$ be the continuous probability function of θ. The following assumption ensures that the government is unable to rule out some realizations of θ.

Assumption. $\phi(\theta) > 0$ for all $\theta \in [\theta_0, \theta_1]$ (H11).

The total level of education is defined as the sum of every individual's education:[7]

$$E = \int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} e(\theta, Y) \phi(\theta) h(Y) \, dY d\theta. \tag{3}$$

Substituting (1) and (2) in the utility function, it can be written as

$$U = u(Y - t - ke) + y(\theta, e; E) + t.$$

3 The Laissez-Faire Equilibrium

The imperfection of educational credit markets was studied, among others, by Becker (1964) and Schultz (1963). It is usually argued that investment in human capital is risky, nondiversifiable, and hard to collateralize, implying that private

[6] See Blaug (1965, pp. 234-241), Cohn (1979) or Lucas (1988) for discussions on the presence of human-capital externalities.

[7] This specification implies the same amount of externality being produced by any unit of education (i.e., the amount of externality caused by a year in high school is the same as preparing for the PhD). However, as will become clear when we present the decentralized scheme, the main results do not depend on such an assumption.

credit markets may fail to finance education. In this economy, credit markets are imperfect in the sense that individuals cannot borrow to finance education.[8]

The parent's problem consists of determining the optimal amount of consumption, and education and bequest left to her daughter. Hence, the parent's problem is

$$\max_{\{e,t\}} u(Y-t-ke)+y(\theta,e;E)+t \qquad s.t.\ t\geq 0.$$

Definition 1 *A competitive equilibrium is a profile of consumption, education, and bequests* $\{c(\theta,Y),e(\theta,Y),t(\theta,Y);\ \theta\in[\theta_0,\theta_1],\ Y\in\Gamma\}$ *and an educational level* E *such that:*

1. $\{c(\theta,Y),e(\theta,Y),t(\theta,Y)\}$ *solve the parent's problem for each* $(\theta,Y)\in[\theta_0,\theta_1]\times\Gamma$ *given the educational level* E*, and*
2. *the educational level is the sum of every individual's education:* $E=\int_{\theta_0}^{\theta_1}\int_{Y\in\Gamma}e(\theta,Y)\,\phi(\theta)\,h(Y)\,dY\,d\theta$.

As usual, the parent will choose the amount of education left to the daughter such that its marginal cost $ku'(c)$ equals its marginal benefit $y_e(\theta,e;E)$. If bequests are positive, investments on education should pay the interest rate (normalized to 1). However, if she does not leave bequests, returns on education should be at least as high as the interest rate.

Define $e^u(\theta,Y,k,E)$ and $e^c(\theta,Y,k,E)$ implicitly by the relations

$$k=y_e(\theta,e^u;E),\quad ku'(Y-ke^c)=y_e(\theta,e^c;E),$$

where the letters u and c stand for unconstrained and constrained, respectively.[9] Solving the household's problem we obtain the following proposition:

Proposition 1 *The laissez-faire competitive equilibrium is* $\{c(\theta,Y),e(\theta,Y),t(\theta,Y);\ \theta\in[\theta_0,\theta_1],\ Y\in\Gamma\}$*, such that*

$$\begin{aligned} e(\theta,Y)&=\max\{e^u(\theta,Y,k,E);e^c(\theta,Y,k,E)\},\\ t(\theta,Y)&=\max\{Y-c^*-ke(\theta,Y,k,E);0\},\\ c(\theta,Y)&=\min\{c^*;Y-ke(\theta,Y,k,E)\},\end{aligned}$$

and $E=\int_{\theta_0}^{\theta_1}\int_{Y\in\Gamma}e(\theta,Y)\,\phi(\theta)\,h(Y)\,dY\,d\theta$.

Proof. The first-order conditions to the household's problem (necessary and sufficient) are

$$\begin{aligned} ku'(Y-t-ke)&=y_e(\theta,e;E),\\ u'(Y-t-ke)&=1+\mu,\\ \min\{t,\mu\}&=0.\end{aligned}$$

[8] This is a usual assumption in education and child-labour models. See, for example, Baland and Robinson (2000) or Ranjan (2001).

[9] The existence and uniqueness of e^u and e^c is demonstrated in the Appendix.

where μ is the Kuhn-Tucker multiplier. If $\mu = 0$, we say that the solution to the problem is unconstrained and it follows that:

$$\begin{aligned} y_e(\theta, e^u; E) &= k, \\ Y - c^* - ke^u &= t^u. \end{aligned}$$

If $\mu > 0$, we say that the solution to the problem is constrained and it follows that:

$$\begin{aligned} y_e(\theta, e^c; E) &= ku'(Y - ke^c) > k \therefore e^c < e^u, \\ u'(Y - ke^c) &> 1 \therefore c^c < c^*. \end{aligned}$$

■

Remark 1 *As can be seen in the proof above, if* $Y < c^* + ke^u(\theta, Y)$*, the parent's decisions are constrained since she would prefer to leave negative bequests but is not allowed to. So, she partially reduces her consumption and partially reduces her daughter's education, and leaves no bequests. Since education is increasing in* θ*, households with sufficiently bright children (high* θ*) or low wealth* Y *are constrained.*[10] *Thus the laissez-faire equilibrium is characterized by inequality of opportunities in the sense that individuals with the same ability receive different amounts of education. Moreover, the marginal productivity of children from constrained parents is higher than those of children from unconstrained parents.*

The following definition will be useful when we allow for voluntary participation in public education. Let $P(\theta, Y, E)$ be the utility obtained in the laissez-faire equilibrium by an individual with wealth Y and whose child's ability parameter is θ. More precisely,

$$P(\theta, Y, E) \equiv \left\{ \begin{array}{c} \max_{\{e,t\}} \ u(Y - t - ke) + y(\theta, e; E) + t \\ s.t. \ t \geq 0 \end{array} \right\}. \tag{4}$$

4 The Government-Intervention Solution

4.1 The First-Best Equilibrium

As in most public-finance literature, we take a government that maximizes the unweighted sum of each parent's utilities.

As usual, we shall refer to the outcome chosen by the government if ability were observable as the first-best equilibrium:

[10] Applying the implicit function theorem, we get:

$$\begin{aligned} \frac{\partial e^u}{\partial \theta} &= -\frac{y_{e\theta}(\theta, e^u; E)}{y_{ee}(\theta, e^u; E)} > 0, \\ \frac{\partial e^c}{\partial \theta} &= -\frac{y_{e\theta}(\theta, e^c; E)}{y_{ee}(\theta, e^u; E) + k^2 u''(Y - ke^c)} > 0. \end{aligned}$$

Definition 2 *A first-best equilibrium is a profile of education and bequests* $\{e(\theta, Y),$ $t(\theta, Y);\ \theta \in [\theta_0, \theta_1],\ Y \in \Gamma\}$ *solving*

$$\max_{\{e(\theta,Y),t(\theta,Y),E\}} \int_{\theta_0}^{\theta_1} \int_{Y\in\Gamma} \left[\begin{array}{c} u(Y - t(\theta, Y) - ke(\theta, Y)) \\ +y(\theta, e(\theta, Y); E) + t(\theta, Y) \end{array} \right] \phi(\theta) h(Y)\, dY d\theta$$

$$s.t. \quad E = \int_{\theta_0}^{\theta_1} \int_{Y\in\Gamma} e(\theta, Y) \phi(\theta) h(Y)\, dY d\theta.$$

For notational convenience we shall define the expectations operator $\bar{E}[\cdot]$ as

$$\bar{E}[e] \equiv \int_{\theta_0}^{\theta_1} \int_{Y\in\Gamma} e(\theta, Y) \phi(\theta) h(Y)\, dY d\theta,$$

where $e = \{e(\theta, Y);\ \theta \in [\theta_0, \theta_1],\ Y \in \Gamma\}$.

Notice that the marginal benefit of education consists of the private marginal return of education y_e and the social return of education $\bar{E}[y_E]$. Hence, the first-best amount of education should be such that the marginal benefit of education equals its marginal cost k. Define e^* implicitly by the relation

$$k = y_e\left(\theta, e^*(\theta, Y); \bar{E}[e^*]\right) + \bar{E}\left[y_E\left(\theta, e^*(\theta, Y), \bar{E}[e^*]\right)\right]. \qquad (5)$$

Let t^* be defined as $t^*(\theta, Y) = Y - ke^*(\theta, Y) - c^*$.

Assumption. $\bar{E}[t^*] \geq 0$. (H12)

The assumption above guarantees that there are enough resources so that e^* and c^* are feasible in a perfect-information economy.[11]

Proposition 2 *The first-best allocations are*

$$\{c^*, e^*(\theta, Y), t^*(\theta, Y); \theta \in [\theta_0, \theta_1], Y \in \Gamma\}.$$

Proof. Introducing the auxiliary variable $S(\theta)$, (3) can be rewritten as

$$\begin{aligned} \dot{S}(\theta) &= \int_{Y\in\Gamma} h(Y) e(\theta, Y) \phi(\theta)\, dY, \\ S(\theta_0) &= 0,\ S(\theta_1) = E. \end{aligned}$$

The optimal policy offered to an individual with wealth Y must solve the following Hamiltonian:

$$H = \int_{Y\in\Gamma} \left[\begin{array}{c} u(Y - t(\theta, Y) - ke(\theta, Y)) \\ +y(\theta, e(\theta, Y); E) + t(\theta, Y) \end{array} \right] \phi(\theta) h(Y)\, dY + \mu(\theta) \int_{Y\in\Gamma} h(Y) e(\theta, Y) \phi(\theta)\, dY,$$

where t and e are control variables and S is a state variable. The first-order conditions are

$$[-u'(Y - t(\theta, Y) - ke(\theta, Y)) + 1]\, h(Y) \phi(\theta) = 0,$$
$$[-k\ u'(Y - t(\theta, Y) - ke(\theta, Y)) + y_e(\theta, e(\theta, Y), E) + \mu(\theta)]\, h(Y) \phi(\theta) = 0,$$
$$\mu(\theta) = \mu \text{ constant}.$$

[11] The existence and uniqueness of e^* is demonstrated in the Appendix.

Let W be the welfare function. Then, as $\left.\frac{\partial W}{\partial E}\right|_{e=e^*,E=E^*} = \mu(\theta_1) = \mu$, it follows that

$$\bar{E}\left[y_E(\theta, e; E)\right] = \mu.$$

Substituting in the first-order conditions, we get the result above. ■

Remark 2 *As the first-best amount of education $e^*(\theta, Y)$ is independent of the household's wealth Y, it follows that the optimal educational policy is characterized by equality of opportunities. Since efficiency requires that marginal productivity of education must be equalized for all individuals, the amount of education received by an individual should depend only on her ability. Therefore, we shall denote $e^*(\theta, Y)$ as $e^*(\theta)$ in order to emphasize that it does not depend on Y. Notice that as c^* is independent of Y, the optimal consumption level is also independent of wealth.*

Remark 3 *Because marginal productivity of education is increasing in ability, it follows that education provided in the first-best solution is also increasing in ability (i.e., the first-best equilibrium is input-regressive).*[12] *Moreover, equality of opportunities in education implies that the first-best equilibrium is output-regressive (i.e., individuals with higher ability obtain more utility than lower-ability individuals). Hence, the first-best equilibrium is characterized by an inequality of outcomes.*

Remark 4 *Notice that the presence of positive externalities implies an inefficiently low amount of education provided in the laissez-faire equilibrium even for unconstrained households (since y is strictly concave in e).*

4.2 The Second-Best Equilibrium

Consider a government that can offer a tax schedule and an education schedule. A tax schedule consists of an income tax $\tau(Y)$. An education schedule consists of an offer of education $e(\theta, Y)$, an up-front fee $f(\theta, Y)$ and a deferred payment $m(\theta, Y)$. Since $f(\theta, Y)$ and $m(\theta, Y)$ may be positive or negative, the government is able to offer loans to students.[13] We assume that the parent's wealth Y is observable but the daughter's ability θ is private information.

With no loss of generality, we can normalize each household's bequests to zero. In that case, all bequests are left through up-front fees and deferred payments. The household's budget constraint is

$$Y = c(\theta, Y) + \tau(Y) + f(\theta, Y).$$

Substituting in the mother's utility function, it can be written as

$$U(\theta, Y) = u(Y - \tau(Y) - f(\theta, Y)) + y(\theta, e(\theta, Y); E) - m(\theta, Y). \quad (6)$$

[12] Applying the implicit function theorem, we get $\frac{\partial e^*(\theta,Y)}{\partial \theta} = -\frac{y_{e\theta}}{y_{ee}} > 0$.

[13] By allowing the government to charge deferred payments, we focus on children above some minimum age. As Becker and Murphy (1988) argue, young children usually cannot be a party to this type of contract.

From the revelation principle, the search for an optimal educational policy can be restricted to the class of incentive-compatible mechanisms with no loss of generality. The following lemma, whose proof is presented in the Appendix, allows us to substitute the incentive-compatibility constraint for the local first- and second-order conditions.

Lemma 1 *A C^2 by parts policy $\{\tau(Y), f(\theta, Y), m(\theta, Y), e(\theta, Y);\ \theta \in [\theta_0, \theta_1],\ Y \in \Gamma\}$ is incentive-compatible if, and only if, it satisfies*

$$U_\theta(\theta, Y) = y_\theta(\theta, e(\theta, Y); E), \tag{7}$$

$$e_\theta(\theta, Y) \geq 0, \tag{8}$$

for all $\theta \in [\theta_0, \theta_1],\ Y \in \Gamma$.

We also assume that individuals are not forbidden to purchase education in the private sector. Hence, they will only join the educational program when their utility exceeds the utility obtained if they purchase education privately. Then, the household's utility must satisfy

$$U(\theta, Y) \geq P(\theta, Y - \tau(Y), E), \quad \forall Y, \forall \theta. \tag{9}$$

4.2.1 The Government Budget Constraint

Up to this point, our model is similar to De Fraja (2002). The distinct feature is that we will enable the government to transfer resources between generations. In the model presented at De Fraja (2002), the budget constraint is imposed to be balanced with each generation in every period. More specifically, the budget constraint is

$$\int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) [\tau(Y) + f(\theta, Y)] \phi(\theta) \, dY d\theta \geq \int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) [ke(\theta, Y)] \phi(\theta) \, dY d\theta,$$

$$\int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} m(\theta, Y) \phi(\theta) \, dY d\theta \geq 0,$$

where the first equation states that the total amount of income taxes and up-front fees is used to finance the educational expenses while the second equation states that the total amount of deferred payments is non-negative.

In each period a mother with ability θ and wealth Y pays $\tau(Y) + f(\theta, Y)$ as up-front taxes and $m(\theta, Y)$ as deferred payments (due to the education received in her childhood) and receives $ke(\theta, Y)$ as education subsidies. However, there is no clear reason why the government would not be able to use the revenue from deferred payments to finance its educational expenses. If the government is able to finance its current educational expenses through deferred payments, then the budget constraint is

$$\int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) [\tau(Y) + f(\theta, Y) + m(\theta, Y)] \phi(\theta) \, dY d\theta \geq$$

$$\int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y)\left[ke(\theta, Y)\right] \phi(\theta)\, dY d\theta. \tag{10}$$

Equation (10) states that the net tax revenue (L.H.S.) is enough to finance the educational expenses (R.H.S.).

Remark 5 *As deferred payments and education are the only channels of transferring wealth between generations, it follows that imposing that the aggregate amount of deferred payments be non-negative is equivalent to imposing that education is the only means of transferring wealth between generations. However, as public educational systems are usually financed through taxes paid by other generations, it may seem reasonable to allow taxes to be used in order to transfer resources between different generations.*

As usual, we shall refer to the optimal contract chosen by the government when ability is not observable as the second-best equilibrium:

Definition 3 *A second-best equilibrium is a policy* $\{\tau(Y), f(\theta, Y), m(\theta, Y), e(\theta, Y);\ \theta \in [\theta_0, \theta_1],\ Y \in \Gamma\}$ *solving*

$$\max_{\{e,\tau,f,m,E\}} \int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) \begin{bmatrix} u(Y - \tau(Y) - f(\theta, Y)) \\ +y(\theta, e(\theta, Y); E) - m(\theta, Y) \end{bmatrix} \phi(\theta)\, dY d\theta$$
$$\text{s.t.}(3),(6),(7),(8),(9),(10).$$

The following proposition, whose proof is presented in the Appendix, is the main result of this article. It ensures that the government is able to implement the efficient level of education and consumption in an economy with private information and where an individual may choose not to join the public educational system. Moreover, since this implementation does not require any additional resources, it achieves first-best welfare.

Proposition 3 *The optimal educational policy implements the first-best amount of education and consumption* $\{e^*(\theta), c^*;\ \theta \in [\theta_0, \theta_1],\ Y \in \Gamma\}$ *and achieves first-best welfare.*

The basic intuition behind this result is that when the government raises the up-front fee uniformly and decreases the income tax in the same amount, the indirect utility of an individual participating in the proposed scheme remains constant, whereas the indirect utility of an individual who purchases education privately decreases. Hence, the participation constraint can be implemented at no cost. Moreover, since the individuals are risk-neutral, any redistribution of wealth does not change the total welfare.

Remark 6 *As shown in Proposition 3, when transfers between generations are allowed, the optimal educational policy provides equality of opportunities in education (since* $e^*(\theta)$ *does not depend on* Y*). Furthermore, as was shown in the* *previous section, the efficient amount of education is higher than the amount provided in the laissez-faire equilibrium. Hence, contrary to the results obtained by De Fraja (2002), the amount of education and consumption does not depend on each parent's wealth.*

In the next section, we show that the optimal educational policy can be implemented through Pigouvian taxes and public provision of credit. This implementation is desirable due to its simplicity and informational advantage.

4.3 Implementation through Pigouvian taxes

Since Pigou (1938), economists know that efficiency in an externality-generating activity can be reached through the imposition of Pigouvian taxes.[14] As Carlton and Loury (1980) show, efficiency may require an additional lump-sum tax-subsidy scheme. In this section, we show that the optimal mechanism can be decentralized through appropriate Pigouvian taxes and the provision of credit at the market interest rate. Moreover, the decentralized scheme does not require knowledge of household's wealth.

We will restrict the space of contracts to those consisting of lump-sum taxes, a linear up-front education fee, and a deferred payment. Formally, let $\tau(Y)$, $f(\theta, Y)$ and $m(\theta, Y)$ be the income tax, up-front fee and deferred payments as defined in the last section. Define $t(\theta, Y)$ and $\hat{k}(\theta, Y)$ as

$$\begin{aligned} t(\theta, Y) &= -m(\theta, Y), \\ \hat{k}(\theta, Y) &= \frac{f(\theta, Y) - t(\theta, Y)}{e(\theta, Y)}. \end{aligned}$$

In general, $\hat{k}$ could depend on θ and Y and τ could depend on Y. However, as we show below, they are both constant for all θ and Y under the optimal policy. Hence, a contract consists of a lump-sum tax τ, a linear education up-front fee $t(\theta, Y) + \hat{k}e(\theta, Y)$ and a deferred payment $-t(\theta, Y)$ (which is a subset of the class of contracts considered previously). This mechanism can be alternatively interpreted as a lump-sum tax τ, a loan $-t(\theta, Y)$ and an up-front fee $\hat{k}e(\theta, Y)$.

Substituting the definitions of $\hat{k}$ and t in the government budget constraint, it can be written as

$$\int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) \left[\left(k - \hat{k} \right) e(\theta, Y) - \tau \right] \phi(\theta) \, dY d\theta \leq 0. \tag{11}$$

In each period, the government pays $\left(k - \hat{k} \right)$ as a subsidy on each unit of education and receives τ as a lump-sum tax. The government also loans $\bar{E}[-t]$ in the first period and receives it in the next period. Since the market interest rate is normalized to 1, $\bar{E}[-t]$ may take any value because it is always repaid in the following period. Thus, as usual, the budget constraint simply states that the total expenses should not exceed the total revenues of the government.

[14] See Baumol (1972) and Kopczuk (2003).

Substituting the definitions of $\hat{k}$ and t in the parent's budget constraint, it follows that the total amount of consumption, loans repaid, and taxes must be equal to the household's wealth:

$$Y = c + t(\theta, Y) + \tau + \hat{k}e(\theta, Y). \tag{12}$$

Hence, we can write the mother's problem as:

$$\max_{\{e(\theta,Y),t(\theta,Y)\}} u\left(Y - t(\theta, Y) - \tau - \hat{k}e(\theta, Y)\right) + y(\theta, e(\theta, Y); E) + t(\theta, Y) \tag{13}$$
$$s.t.\ E = \bar{E}[e]$$

As there are no restrictions on t, the solution must be such that the marginal utility of consumption is equal to the marginal utility of wealth left to the daughter. Hence, each parent must be consuming c^*. Moreover, the marginal benefit of education y_e must be equal to its marginal cost $\hat{k}$. This result is stated formally in the following lemma:[15]

Lemma 2 *The solution to the mother's problem (13) is* $\{c^P(\theta, Y), e^P(\theta, Y), t^P(\theta, Y);\ \theta \in [\theta_0, \theta_1],\ Y \in \Gamma\}$, *such that:*

$$c^P(\theta, Y) = c^*,$$
$$\hat{k} = y_e\left(\theta, e^P(\theta, Y); E\right), \tag{15}$$
$$t^P(\theta, Y) = Y - c^* - \tau - \hat{k}e^P(\theta, Y).$$

Proof. The result follows from the first-order conditions of the household's problem (13). ■

Now, we are ready to show that the first-best solution can be reached through a suitable choice of τ and $\hat{k}$. As in Carlton and Loury (1980), the efficient allocation can be reached in this context through Pigouvian subsidies and a lump-sum tax. This result can be seen as an application of the so-called 'Principle of Targeting' according to which externalities should be corrected by targeting its source directly.

The price of education $\hat{k}$ is chosen in order to internalize for the educational externalities. Hence, it must be equal to the private cost of education k minus the educational externalities $\bar{E}[y_E]$. The lump-sum tax is set in order to cover the expenses from the subsidies. Therefore, it must be equal to the average subsidy $\bar{E}\left[e^P\right]\bar{E}[y_E]$.

Proposition 4 *There exists a second-best equilibrium where the price of education* $\hat{k}$ *and the tax* τ *are both constant in* (θ, Y). *Moreover, this equilibrium implements first-best welfare.*

[15] The existence of e^P is demonstrated in the Appendix.

Proof. Set $\hat{k}$ as

$$\hat{k} = k - \bar{E}\left[y_E(\theta, e^*, \bar{E}[e^*])\right]. \tag{16}$$

Substituting in the first-order conditions of the household's problem (15), we get $e^P(\theta, Y) = e^*(\theta, Y)$.

Set τ as

$$\tau = \bar{E}[e^*] \;\; \bar{E}\left[y_E(\theta, e^*, \bar{E}[e^*])\right]. \tag{17}$$

Then, it follows that

$$\bar{E}\left[t^P\right] = \bar{E}\left[Y - \hat{k}e^* - c^* - \tau\right] = \bar{E}\left[Y - ke^* - c^*\right] = \bar{E}\left[t^*\right].$$

Hence, as c^*, $e^*(\theta, Y)$, and $\bar{E}[t^*]$ are the same as in the first-best solution (and utility is linear in t), first-best welfare is achieved.

From equation (16), it follows that

$$\left(k - \hat{k}\right) e^*(\theta, Y) = \bar{E}\left[y_E(\theta, e^*, \bar{E}[e^*])\right] \;\; e^*(\theta, Y).$$

Applying $\bar{E}$ to both sides of the above expression yields

$$\bar{E}\left[\left(k - \hat{k}\right) e^*\right] = \bar{E}\left[y_E(\theta, e^*, \bar{E}[e^*])\right] \;\; \bar{E}[e^*]. \tag{18}$$

Hence, equations (17) and (18) imply that $\bar{E}\left[\left(k - \hat{k}\right) e^* - \tau\right] = 0$. It follows that the government's budget constraint (11) is satisfied. ■

Remark 7 *As $\bar{E}\left[t^P\right] = \bar{E}[t^*] > 0$, the government transfers resources from older individuals to younger individuals (who repay when older). Thus, this mechanism does not satisfy the government budget constraint imposed in De Fraja (2002) which states that deferred payments can not be used to finance other expenses (i.e., $\bar{E}[t] \leq 0$).*

Define the household's financial contribution as

$$z(\theta, Y) \equiv \tau + \hat{k}e^*(\theta).$$

As education is independent of wealth, it is clear that an individual's financial contribution is independent of her income. Moreover, $z(\theta)$ is strictly increasing in ability since $\hat{k} > 0$. Therefore, households with brighter children contribute more than households with less bright children. These results differ from De Fraja (2002), where households with higher incomes contribute less than those with lower incomes and households with less bright children contribute more than those with brighter children.

5 The economy with default

In the model presented, the returns to education are deterministic. In reality, however, individuals can neither be sure about finishing their education successfully nor about their future returns after a successful conclusion. Hence, educational returns display a very high variation since students may not graduate or not find a job.[16]

Usually, credit provision in an uncertain environment requires the provision of collateral. However, unlike other types of investments, antislavery laws preventing the repossession of human capital precludes the use of human capital as collateral.

In this section, we extend the model presented before to incorporate the possibility of default. We make the assumption that unemployed individuals cannot be charged for their debts. The probability of being unemployed depends on the ability-type of the individual and the parent's wealth.

Usually, the existence of default would result in the incidence of adverse selection since low-ability individuals would be associated with higher probability of default and, thus, might prefer to lie about their ability. However, we shall show that the basic results obtained in the model with no default are still valid in this case since the incidence of adverse selection is totally mitigated in the optimal educational policy.

Let $\psi : [\theta_0, \theta_1] \times \Gamma \to [0,1]$ be the proportion of individuals with type θ and parent's wealth Y who are able to repay the deferred payments $m(\theta, Y) \geq 0$. We assume that $\psi_\theta(\theta, Y) > 0$. Although no restrictions on the dependence of $\psi(\theta, Y)$ on Y are needed for our results, it may seem reasonable to assume that ψ is increasing in Y.

Given the (τ, f, m, e, t), the utility of a mother with wealth Y and whose daughter has ability parameter θ is:

$$U(\theta, Y) = u(Y - \tau(Y) - f(\theta, Y)) + \\ y(\theta, e(\theta, Y); E) - \psi(\theta, Y) m(\theta, Y) + t(\theta, Y), \quad (19)$$

where $t(\theta, Y) \geq 0$ is the amount of bequest and $m(\theta, Y) \geq 0$ is the amount of deferred payments.

Define $\hat{U} : [\theta_0, \theta_1]^2 \times \Gamma \to \mathbb{R}$ as the utility received by an individual with wealth Y and with a type-θ child who gets a contract designed for an individual with a type-$\tilde{\theta}$ child:

$$\hat{U}\left(\hat{\theta}, \theta, Y\right) = u\left(Y - \tau(Y) - f\left(\hat{\theta}, Y\right)\right) + \\ y\left(\theta, e\left(\hat{\theta}, Y\right); E\right) - \psi(\theta, Y) m\left(\hat{\theta}, Y\right) + t\left(\hat{\theta}, Y\right).$$

[16] The high variation of educational returns was originally pointed out in Becker (1964, pp. 104). For a recent study on this issue, see Miller and Volker (1993).

Then, the incentive-compatibility constraint is

$$\hat{U}(\theta,\theta,Y) \geq \hat{U}\left(\hat{\theta},\theta,Y\right), \ \forall\hat{\theta},\theta.$$

As in Lemma 1, the incentive-compatibility constraint can be written as the local first- and second-order conditions.

Lemma 3 *A C^2 by parts policy $\{\tau(Y), f(\theta,Y), m(\theta,Y), e(\theta,Y);\ \theta\in[\theta_0,\theta_1],\ Y\in\Gamma\}$ is incentive-compatible if and only if it satisfies*

$$U_\theta(\theta,Y) = y_\theta(\theta, e(\theta,Y); E) - \psi_\theta(\theta,Y)\, m(\theta,Y), \tag{20}$$

$$y_{e\theta}(\theta, e(\theta,Y); E)\, e_\theta(\theta,Y) - \psi_\theta(\theta,Y)\, m_\theta(\theta,Y) \geq 0. \tag{21}$$

The government's budget constraint states that the educational expenditures must be financed through taxes:

$$\bar{E}\left[\tau(Y) + f(\theta,Y) + \psi(\theta,Y)\, m(\theta,Y) - t(\theta,Y)\right] \geq \bar{E}\left[ke(\theta,Y)\right]. \tag{22}$$

Then, the government faces the following problem:

$$\begin{array}{c} \max\limits_{\{e,\tau,f,m,E\}} \int_{\theta_0}^{\theta_1}\int_{Y\in\Gamma} h(Y)\, U(\theta,Y)\,\phi(\theta)\, dY d\theta \\ s.t.\ (3),\ (9),\ (19),\ (20),\ (21),\ (22), \\ m(\theta,Y)\geq 0, \\ t(\theta,Y)\geq 0. \end{array}$$

Solving this problem we get the following proposition, whose proof is presented in the Appendix:

Proposition 5 *The optimal educational policy in the economy with default implements the first-best amount of education and consumption $\{e^*(\theta,Y), c^*;\ \theta\in[\theta_0,\theta_1],\ Y\in\Gamma\}$ and achieves first-best welfare.*

The basic intuition behind this result is that when each parent is risk-neutral in the wealth left to her daughter, it is indifferent between a certain payment and a lottery with the same expected value. Hence, the certain deferred payment in the environment with no default can be substituted by a lottery where the payment occurs only when the daughter's labor income allows her to.

Remark 8 *If ψ is interpreted as the probability that the labor income of a type-θ individual is higher than a threshold Y^*, then the optimal educational policy can also be implemented through an income-contingent type of payment since the deferred payment is charged only if the realization of income exceeds this threshold.*[17]

[17] See Krueger and Bowen (1993) and Barr (1991) for discussions on income-contingent payments as a means of financing education.

6 Conclusion

In this paper, we show that when the government is able to transfer resources between generations, the optimal educational policy achieves the same amount of welfare that could be reached if ability were observable (first-best efficiency). Moreover, the implementation of the optimal policy can be decentralized through Pigouvian taxes and student loans. An advantage of decentralization is that it is less informationally demanding: the government may not know each household's wealth or ability (it is sufficient to know the optimal level of externality and the social marginal benefit it causes).

When returns to education are random, the non-transferability of human capital implies the emergence of default. In this case, the optimal educational policy can be implemented by a type of income-contingent loan and still achieves first-best efficiency.

As first-best efficiency requires that marginal productivity of education be equalized across individuals, it follows that the amount of education received by a child does not dependent on his/her parent's wealth. Therefore, *equality of opportunities in education is provided.* Furthermore, since the optimal amount of financial contribution does not depend on parental income and is increasing in the ability of the child, *the optimal educational policy is not regressive* (i.e., wealthier households do not contribute less than poorer households).

The optimal educational policy is input-regressive (in the sense of Arrow, 1971) since more able individuals receive higher education than those less able. Moreover, as more able individuals attain a higher utility, there is an inequality of outcomes (i.e., the policy is output-regressive).

The laissez-faire equilibrium was inefficient due to imperfect credit markets and educational externalities. We have shown that the government should provide student loans in order to correct the credit-market inefficiency. Governmental provision of credit is probably the educational policy most suggested.[18] According to Becker (1991, p.188):

> 'Public (or private) policies that improve access to the capital markets by poorer families - perhaps a loan program to finance education (...) - would increase the efficiency of society's investments in human capital while equalizing opportunity and reducing inequality.'

By not internalizing the effects that education bears on the rest of the economy, the amount of education that each household provides in the laissez-faire equilibrium is inefficiently low. We show that the first-best solution can be achieved through Pigouvian taxes. In this context, the appropriate Pigouvian taxes are educational subsidies that induce households to internalize for the (positive) externalities caused by education.[19]

[18] See, for example Barr (1991, 1993), Becker (1991), and Krueger and Bowen (1993). As Eden (1994) remarks: "Government backed loans can mitigate capital market imperfections and most economists will favor this type of intervention."

[19] Friedman (1955, pp.124-125) advocated for a scheme similar to the Pigouvian taxes pro-

A Appendix

A.1 Proof of Lemma 1:

Define $\hat{U} : [\theta_0, \theta_1]^2 \times \Gamma \to \mathbb{R}$ as the utility received by a type θ individual with wealth Y who gets a contract designed for a type $\tilde{\theta}$ individual:

$$\hat{U}\left(\tilde{\theta}, \theta, Y\right) = u\left(Y - \tau(Y) - f\left(\tilde{\theta}, Y\right)\right) + y\left(\theta, e\left(\tilde{\theta}, Y\right); E\right) - m\left(\tilde{\theta}, Y\right).$$

In order to be incentive-compatible, each individual must prefer to announce his own type. Hence, the following first- and second-order conditions must be satisfied for almost all θ:

$$\left.\frac{\partial \hat{U}\left(\tilde{\theta}, \theta, Y\right)}{\partial \tilde{\theta}}\right|_{\tilde{\theta}=\theta} = 0$$

$$\left.\frac{\partial \hat{U}^2\left(\tilde{\theta}, \theta, Y\right)}{\partial \tilde{\theta}^2}\right|_{\tilde{\theta}=\theta} \leq 0$$

The first-order condition yields, for almost all θ,

$$-u'(Y - \tau(Y) - f(\theta, Y)) f_\theta(\theta, Y) + y_e(\theta, e(\theta, Y); E) e_\theta(\theta, Y) - m_\theta(\theta, Y) = 0.$$

Differentiating the first-order condition, we get

$$\begin{aligned} -u'(Y - \tau(Y) - f(\theta, Y)) f_{\theta\theta}(\theta, Y) + u''(Y - \tau(Y) - f(\theta, Y)) [f_\theta(\theta, Y)]^2 \\ + y_{e\theta}(\theta, e(\theta, Y); E) e_\theta(\theta, Y) + y_{ee}(\theta, e(\theta, Y); E) [e_\theta(\theta, Y)]^2 \\ + y_e(\theta, e(\theta, Y); E) e_{\theta\theta}(\theta, Y) - m_{\theta\theta}(\theta, Y) = 0. \end{aligned} \tag{23}$$

The second-order condition yields, for almost all θ,

$$\begin{aligned} -u'(Y - \tau(Y) - f(\theta, Y)) f_{\theta\theta}(\theta, Y) + u''(Y - \tau(Y) - f(\theta, Y)) [f_\theta(\theta, Y)]^2 \\ + y_{ee}(\theta, e(\theta, Y); E) [e_\theta(\theta, Y)]^2 + y_e(\theta, e(\theta, Y); E) e_{\theta\theta}(\theta, Y) - m_{\theta\theta}(\theta, Y) \leq 0. \end{aligned} \tag{24}$$

Substituting (23) in equation (24), we obtain

$$y_{e\theta}(\theta, e(\theta, Y); E) e_\theta(\theta, Y) \geq 0.$$

As $y_{e\theta}(\theta, e(\theta, Y); E) > 0$, this equation is equivalent to the monotonicity condition $e_\theta(\theta, Y) \geq 0$.

posed here. He argued that since buyers of education generate external benefits for those not purchasing education, the government should subsidize those purchasing education and tax those who are not.

Differentiating equation (6), yields

$$\begin{aligned} U_\theta(\theta, Y) &= -u'(Y - \tau(Y) - f(\theta, Y)) f_\theta(\theta, Y) + y_\theta(\theta, e(\theta, Y); E) \\ &\quad + y_e(\theta, e(\theta, Y); E) e_\theta(\theta, Y) - m_\theta(\theta, Y). \end{aligned}$$

Substituting the first-order condition in this expression, we get

$$U_\theta(\theta, Y) = y_\theta(\theta, e(\theta, Y); E).$$

This proves the necessity of (7) and (8).

To prove the sufficiency of (7) and (8), assume that a type θ strictly prefers to announce $\tilde{\theta} \neq \theta$:

$$\hat{U}\left(\tilde{\theta}, \theta, Y\right) > \hat{U}(\theta, \theta, Y).$$

This equation can be rewritten as $\int_\theta^{\tilde{\theta}} \hat{U}_1(x, \theta, Y)\, dx > 0$, where $\hat{U}_1(x, \theta, Y) \equiv \frac{\partial \hat{U}(\tilde{\theta}, \theta, Y)}{\partial \tilde{\theta}}$. As $\hat{U}_1(x, x, Y) = 0$ for almost all x, it follows that

$$\int_\theta^{\tilde{\theta}} \left[\hat{U}_1(x, \theta, Y) - \hat{U}_1(x, x, Y)\right] dx = \int_\theta^{\tilde{\theta}} \int_x^\theta \hat{U}_{12}(x, z, Y)\, dz dx > 0.$$

As $\hat{U}_{12}\left(\tilde{\theta}, \theta, Y\right) = y_{e\theta}\left(\theta, e\left(\tilde{\theta}, Y\right); E\right)\ e_\theta\left(\tilde{\theta}, Y\right) \geq 0$ and x is between θ and $\tilde{\theta}$, this inequality cannot hold. ■

A.2 Proof of Proposition 5:

Introducing the auxiliary variable $S(\theta)$, (3) can be rewritten as

$$\begin{aligned} \dot{S}(\theta) &= \int_{Y \in \Gamma} h(Y) e(\theta, Y) \phi(\theta)\, dY, \qquad (25) \\ S(\theta_0) &= 0,\ S(\theta_1) = E. \end{aligned}$$

Analogously, we introduce the auxiliary variables $x(\theta, Y) = m_\theta(\theta, Y),\ z(\theta, Y) = e_\theta(\theta, Y)$. Let $Y \in \Gamma$ be an arbitrary wealth level. For the moment, we will ignore the monotonicity condition (8). Hence, we will consider the following relaxed problem:

$$\begin{gathered} \max \int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) [u(Y - \tau - f) + y + \tau + f - ke] \phi(\theta)\, dY d\theta \\ s.t.\quad \dot{S}(\theta) = \int_{Y \in \Gamma} he\phi dY, \quad U_\theta = y_\theta - \psi_\theta m, \\ m_\theta = x, \quad e_\theta = z, \quad U = u(Y - \tau - f) + y - \psi m + t, \\ y_{e\theta} z - \psi_\theta x \geq 0, \quad U \geq P(\theta, Y - \tau, E), \quad m \geq 0, \quad t \geq 0, \end{gathered}$$

where the state variables are U, e, m, S, the control variables are f, τ, x, t, z and we omit the dependence on θ, Y and E for notational clarity. Then, the optimal policy offered to an individual with wealth Y must solve the following Hamiltonian:

$$
\begin{aligned}
H = & \int_{Y\in\Gamma} h(Y)\left[u(Y-\tau-f)+y+\tau+f-ke\right]\phi(\theta)\,dY \\
& +\mu^1(\theta)\int_{Y\in\Gamma} h(Y)\,e\phi(\theta)\,dY+\mu^2(\theta,Y)\left[y_\theta-\psi_\theta m\right] \\
& +\mu^3(\theta,Y)\,x+\mu^4(\theta,Y)\,z \\
& -\rho^1(\theta,Y)\left[u(Y-\tau-f)+y-\psi m+t-U\right] \\
& +\lambda^1(\theta,Y)\left[y_{e\theta}z-\psi_\theta x\right]+\lambda^2(\theta,Y)\left[U-P(\theta,Y-\tau,E)\right] \\
& +\lambda^3(\theta,Y)\,m+\lambda^4(\theta,Y)\,t(\theta,Y).
\end{aligned}
$$

From the first order conditions, it follows that:

$$\rho^1(\theta,Y)=h(Y)\,\phi(\theta)\left[1-\frac{1}{u'(Y-\tau-f)}\right], \tag{26}$$

$$\lambda^2(\theta,Y)=0, \text{ for almost all } Y,\theta, \tag{27}$$

$$\mu^3(\theta,Y)=\lambda^1(\theta,Y)\,\psi_\theta, \tag{28}$$

$$\rho^1(\theta,Y)=\lambda^4(\theta,Y), \tag{29}$$

$$\mu^4(\theta,Y)=-\lambda^1(\theta,Y)\,y_{e\theta}, \tag{30}$$

$$\rho^1(\theta,Y)=-\mu^2_\theta(\theta,Y), \tag{31}$$

$$
\begin{aligned}
-\mu^4_\theta(\theta,Y) \quad = \quad & h(Y)\,\phi(\theta)\left[y_e-k+\mu^1(\theta,Y)\right]+\mu^2(\theta,Y)\;\, y_{e\theta} \\
& -\rho^1(\theta,Y)\;\, y_e+\lambda^1(\theta,Y)\;\, z\; y_{ee\theta},
\end{aligned} \tag{32}
$$

$$\mu^1(\theta)=\mu^1 \text{ constant}, \tag{33}$$

$$-\mu^3_\theta(\theta,Y)=-\mu^2(\theta,Y)\,\psi_\theta+\rho^1(\theta,Y)\,\psi+\lambda^3(\theta,Y). \tag{34}$$

$$\min\{\lambda^2; U-P(\theta,Y-\tau,E)\}=0 \tag{35}$$

Substituting (26) in (31),

$$\mu^2(\theta,Y)=\int_{\theta_0}^{\theta} h(Y)\left[\frac{1}{u'(Y-\tau-f)}-1\right]\phi(\theta)\,d\theta$$

Notice that equation (27) implies that $U(\theta,Y)\geq P(\theta,Y-\tau(Y),E)$ is never binding. Hence, as $U(\theta_0)$ and $U(\theta_1)$ are free, the transversality condition implies that

$$\mu^2(\theta_1,Y)=\int_{\theta_0}^{\theta_1} h(Y)\left[\frac{1}{u'(Y-\tau-f)}-1\right]\phi(\theta)\,d\theta=0.$$

Then, as $\lambda^4(\theta, Y) \geq 0$, from equation (29) we get $u'(Y - \tau - f) \geq 1$, for almost all Y, θ.

If $u'(Y - \tau - f) > 1$ for some set with positive measure, then $\mu^2(\theta_1, Y) < 0$ which contradicts the transversality condition. Hence, it follows that

$$Y - \tau(Y) - f(\theta, Y) = c^*, \text{ for almost all } Y, \theta.$$

Therefore, the second-best amount of consumption is the same as the first-best amount and one must (almost) always face a unitary marginal tax. Moreover, from equation (31),

$$\mu^2(\theta_1, Y) = \rho^1(\theta, Y) = 0, \text{ for almost all } Y, \theta. \tag{36}$$

Then, from equation (34), we get

$$\mu_\theta^3(\theta, Y) = -\lambda^3(\theta, Y) \leq 0. \tag{37}$$

As $m(\theta_0, Y)$ and $m(\theta_1, Y)$ are free, the transversality conditions impose that $\mu^3(\theta_0, Y) = \mu^3(\theta_1, Y) = 0$. Hence,

$$\mu^3(\theta, Y) = \lambda^3(\theta, Y) = 0, \text{ for almost all } Y, \theta. \tag{38}$$

Therefore, equations (28) and (3) imply $\lambda^1(\theta, Y) = \mu^4(\theta, Y) = 0$. Then, from equation (32) we get

$$\mu^1 = k - y_e. \tag{39}$$

Lemma 4 *The amount of education solving the relaxed problem above is the same as in the first-best solution. That is,* $e(\theta, Y) = e^*(\theta)$, *for almost all* $\{\theta, Y\} \in [\theta_0, \theta_1] \times \Gamma$.

Proof. Let e^{2b}, E^{2b} be the amounts of education and externalities that solve the second-best problem defined before Proposition 3. As $\frac{\partial W}{\partial E}\Big|_{e=e^{2b}, E=E^{2b}} = \mu^1(\theta_1) = \mu^1$, it follows that

$$\int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) y_E\left(\theta, e^{2b}, E^{2b}\right) \phi(\theta) dY d\theta = \mu^1.$$

Substituting into (39),

$$y_e\left(\theta, e^{2b}; E\right) = k - \int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) y_E\left(\theta, e^{2b}, E^{2b}\right) \phi(\theta) dY d\theta,$$

which is the equation that defines e^*. ■

Equation (4) implies that $P(\theta, 0, E) = 0$ and, from the envelope theorem, $P_Y(\theta, Y - \tau(Y), E) \geq 1$. Moreover, a unitary increase in $\tau(Y)$ and a unitary decrease in $f(\theta, Y)$ leaves $U(\theta, Y)$ constant. Hence, it is always possible to choose $\tau(Y)$ and $f(\theta, Y)$ such that condition (9) is satisfied.

We have to show that the monotonicity condition (8) is satisfied in the relaxed problem considered. But, as we have already shown in remark (3), $e^*(\theta)$ is increasing in θ. Therefore the monotonicity condition $e_\theta(\theta, Y) \geq 0$ is satisfied.

Hence, the amount of education and consumption solving the problem above is the same as in the first-best solution. Since individual utilities are linear in repaid deferred payments $\psi(\theta, Y) m(\theta, Y)$ and bequests $t(\theta, Y)$, it follows that any profile of deferred payments and bequests such that the government's budget constraint is satisfied as an equality achieves the same welfare as the first-best solution. ■

Proposition 3 follows as a corollary since it is a particular case of the economy with default when $\psi(\theta, Y) = 1$ for all θ, Y.

A.3 Existence and uniqueness of equilibria:

The following proposition ensures that the education profiles in the Laissez-Faire equilibrium are well defined.

Proposition 6 *There exists* e^u *and* e^c *such that* $k = y_e(\theta, e^u; E)$, $ku'(Y - ke^c) = y_e(\theta, e^c; E)$. *Moreover,* e^u *and* e^c *are unique.*

Proof. Define $\xi : \mathbb{R}_+ \to \mathbb{R}_+$ as $\xi(e) \equiv y_e(\theta, e; E)$. Then, as ξ is continuous, $\lim_{e\to 0} \xi(e) = +\infty$ and $\lim_{e\to+\infty} \xi(e) = 0$, it follows that there exists e^u such that $\xi(e^u) = k$. Moreover, as $\xi'(e) = y_{ee}(\theta, e; E) < 0$, e^u is unique.

Analogously define $\varphi : \mathbb{R}_+ \to \mathbb{R}$ as $\varphi(e) \equiv y_e(\theta, e; E) - ku'(Y - ke)$. As $Y > 0$, it follows that $\lim_{e\to 0} \varphi(e) = +\infty$. Then, as φ is continuous and $\lim_{e\to \frac{Y}{k}} \varphi(e) = -\infty$, it follows that there exists e^c such that $\varphi(e^c) = 0$. Furthermore, as $\varphi'(e) = y_{ee}(\theta, e; E) + k^2 u''(Y - ke) < 0$, e^c is unique. ■

The same argument establishes the existence of the education profile defined before Lemma 2.

Corollary 1 *There exists a unique* e^P *such that* $\hat{k} = y_e\left(\theta, e^P(\theta, Y); E\right)$.

The following proposition ensures the existence of the first-best level of education (which is also the second-best level of education).

Proposition 7 *There exists a unique* e^* *such that*

$$k = y_e\left(\theta, e^*(\theta, Y); \bar{E}[e^*]\right) + \bar{E}\left[y_E\left(\theta, e^*(\theta, Y), \bar{E}[e^*]\right)\right].$$

Proof. Notice that if e^* exists, it must be constant in Y. Fix an arbitrary $\theta \in [\theta_0, \theta_1]$ and denote $e(-\theta)$ as $\{e\left(\hat{\theta}\right); \hat{\theta} \neq \theta\}$.

Define the function ρ as

$$\rho(e(\theta), e(-\theta), E) \equiv y_e(\theta, e(\theta); E) + \bar{E}[y_E(\theta, e(-\theta), E)] - k.$$

Then, as $\lim_{e\to 0} \rho(e, e(-\theta), E) = +\infty$, $\lim_{e\to +\infty} \rho(e, e(-\theta), E) < 0$ (by H9 and H10) and ρ is continuous, it follows that, for every $e(-\theta)$ and every E, there exists $\tilde{e}(\theta)$ such that $\rho(\tilde{e}(\theta), e(-\theta), E) = 0$. Moreover, the Inada conditions imply that this $\tilde{e}(\theta)$ is unique.

Since $\lim_{e\to +\infty} \rho(e, e(-\theta), E) < 0$ and ρ is continuous, there exists $\bar{e}$ such that, for all $e > \bar{e}$, $\rho(e, e(-\theta), E) < 0$.

Define ϵ and P as $\epsilon \equiv [0, \bar{e}]$ and $P \equiv \{\{e(\theta)\}_{\theta\in[\theta_0,\theta_1]}\}$. Then, $F \equiv \{(E, e) \in \epsilon \times P;\ E = \int_{\theta_0}^{\theta_1} e(\theta)\phi(\theta)\, d\theta\}$ is a compact, convex set in the product topology.

Define the function $T : F \to F$ as $T(E, e) = \left(\tilde{E}, \tilde{e}\right)$, where $\tilde{e} \equiv \{\tilde{e}(\theta) : \theta \in [\theta_0, \theta_1]\}$ and $\tilde{E} \equiv \int_{\theta_0}^{\theta_1} \tilde{e}(\theta)\phi(\theta)\, d\theta$ (from the definition of $\bar{e}$, it follows that $\tilde{E} \in P$).

Then, the Schauder-Tychonoff Theorem implies the existence of a fixed point of T, $\left(\hat{E}, \hat{e}\right)$ (see, Dunford and Schwartz, 1988, p. 456). From the definition of T, this fixed point must satisfy equation (5) and $\hat{E} = \int_{\theta_0}^{\theta_1} \hat{e}(\theta)\phi(\theta)\, d\theta$.

The uniqueness follows from the strict concavity of the first-best problem. ■

References

[1] ARROW, K.J. (1971). "A Utilitarian Approach to the Concept of Equality in Public Expenditures", *Quarterly Journal of Economics*, **85**, 409-415.

[2] BALAND, J.M. and ROBINSON, J.A. (2000). "Is Child Labor Inefficient?", *Journal of Political Economy*, **108**, 663-679.

[3] BAUMOL, W.J. (1972). "On Taxation and the Control of Externalities", *American Economic Review*, **62**, 307-322.

[4] BARR, N. (1990). "Income-Contingent Student Loans: An Idea Whose Time Has Come", In: SHAW, G.K. (ed.), *Economics, Culture and Education: Essays in Honour of Mark Blaug* (Hants: Edward Elgar), 155-170.

[5] BARR, N. (1993). "Alternative Funding Resources for Higher Education", *Economic Journal*, **103**, 718-728.

[6] BECKER, G. S. (1964). *Human Capital: A Theoretical Analysis with Special Reference to Education.* (New York: Columbia University Press).

[7] BECKER, G. S. (1991). "Family Background and the Opportunities of Children", In: BECKER, G. S. *A Treatise on the Family* (Cambridge: Harvard University Press).

[8] BECKER, G. S. and MURPHY, K.M. (1988). "The Family and the State", *Journal of Law and Economics*, **31**, 1-18.

[9] BLAUG, M. (1965). "The Rate of Return on Investment on Education in Great Britain", *Manchester School*, **33**, 205-251.

[10] CARLTON, D. W. and LOURY, G. C. (1980). "The Limitations of Pigouvian Taxes as a Long-Run Remedy for Externalities", *Quarterly Journal of Economics*, **95**, 559-566.

[11] COHN, E. (1979). *The Economics of Education* (Cambridge: Ballinger Publishing Company).

[12] DE BARTOLOME, C. A. M. (1990). "Equilibrium and Inefficiency in a Community Model with Peer Group Effects", *Journal of Political Economy*, **98**, 110-133.

[13] DE FRAJA, G. (2002). "The Design of Optimal Education Policies", *Review of Economic Studies*, **69**, 437-466.

[14] DE FRAJA, G. (2002b). "Equal Opportunities in Education: Market Equilibrium and Public Policy", Mimeo, University of York.

[15] DUNFORD, N. and SCHWARTZ, J. T. (1988). *Linear Operators, Part I: General Theory* (New York: John Wiley and Sons).

[16] EDEN, B. (1994). "How to Subsidize Education: An Analysis of Voucher Systems", Mimeo, University of Haifa.

[17] EPPLE, D. and ROMANO R. E. (1998). "Competition Between Private and Public Schools, Vouchers, and Peer-Group Effects", *American Economic Review*, **88**, 33-62.

[18] FERNANDEZ, R. and ROGERSON, R. (1995). "On the Political Economy of Education Subsidies", *Review of Economic Studies*, **62**, 249-262.

[19] FERNANDEZ, R. and ROGERSON, R. (1996). "Income Distribution, Communities, and the Quality of Public Education", *Quarterly Journal of Economics*, **111**, 135-164.

[20] FERNANDEZ, R. and ROGERSON, R. (1998). "Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform", *American Economic Review*, **88**, 813-833.

[21] FRIEDMAN, M. (1955). "The Role of Government in Education", In: SOLOW, R. (Ed.), *Economics and the Public interest* (New Brunswick: Rutgers University Press).

[22] HARE, P. G. and ULPH, D. T. (1979). "On Education and Distribution", *Journal of Political Economy*, **87**, 193-212.

[23] INMAN, R. P. (1978). "Optimal Fiscal Reform of Metropolitan Schools: Some Simulation Results", *American Economic Review*, **68**, 107-122.

[24] JOHNSON, G. E. (1984). "Subsidies for Higher Education", *Journal of Labor Economics*, **2**, 303–318.

[25] KOPCZUK, W. (2003). "A note on optimal taxation in the presence of externalities", *Economic Letters*, **80**, 81-86.

[26] KOZOL, J. (1991). *Savage Inequalities* (New York: Crown Publishers).

[27] KRUEGER, A. and BOWEN, W. G. (1993). "Policy-Watch: Income-Contingent College Loans", *Journal of Economic Perspectives*, **7**, 193-201.

[28] LOTT, J. R. (1987). "Why Is Education Publicly Provided? A Critical Survey", *Cato Journal*, **7**, 475–501.

[29] LUCAS, R. E. Jr. (1988). "On the Mechanics of Economic Development", *Journal of Monetary Economics*, **22**, 3-42.

[30] MILLER, P. and VOLKER, P. (1993). "Youth Wages, Risk, and Tertiary Finance Arrangements", *Economic Record*, **69**, 20-33.

[31] MIRRLEES, J. A. (1971). "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies*, **38**, 175-208.

[32] NERLOVE, M., RAZIN, A. and SADKA, E. (1988). "A bequest-constrained economy: Welfare analysis", *Journal of Public Economics*, **37**, 203-220.

[33] PELTZMAN, S. (1973). "The Effect of Government Subsidies-in-Kind on Private Expenditures: The Case of Higher Education", *Journal of Political Economy*, **81**, 1–27.

[34] PIGOU, A. C. (1938). *The Economics of Welfare* (London: MacMillan).

[35] PSACHAROPOULOS, G. (1986). *Financing Education in Developing Countries* (Washington D.C.: World Bank).

[36] RANJAN, P. (2001). "Credit constraints and the phenomenon of child labor", *Journal of Development Economics*, **64**, 81–102.

[37] SHULTZ, T. W. (1963). *The Economic Value of Education* (New York: Columbia University Press).

[38] STIGLITZ, J. E. (1974). "The Demand for Education in Public and Private School Systems", *Journal of Public Economics*, **3**, 349-385.

[39] ULPH, D. (1977). "On the Optimal Income Taxation and Educational Expenditure", *Journal of Public Economics*, **8**, 341-356.