Should educational policies be regressive?

Daniel Gottlieb
Humberto Moreira

Outubro de 2003
Should Educational Policies be Regressive?

Daniel Gottlieb, Humberto Moreira*

Graduate School of Economics, Getulio Vargas Foundation,
Praia de Botafogo, 190, sala 1121, CEP: 22253-900,
Rio de Janeiro, RJ, Brazil.

September 29, 2003

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, students loans, redistribution.

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). 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 is known at least since the work of Mirrlees (1971). In the specific case of

---

*E-mail addresses: gottlieb@fgvmail.br and humberto@fgv.br.

1 See, for example, Fernandez and Rogerson (1996), Kozol (1991) or Psacharopoulos (1986).

In the United States, this regressivity is reflected in the large disparity of spending per students 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.
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 irrelevance of the education received by a child on household’s income). 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) follow from 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 exactly what pay-as-you-go social-security systems are: young generations contribute for the benefits of the older generations. As most social-security systems in the world are (at least partially) pay-as-you-go systems, it seems reasonable to assume that governments are able to transfer between generations.

We show that if transfers between generations are allowed, then the optimal educational policy takes a very different form: it achieves first-best welfare 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 observed inequalities must 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.\(^2\)

The remainder of the paper is organized as follows. Section 2 lays out the basic framework of the model. 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

Consider an economy with a continuum of households with measure normalized to 1. Each household consists of a parent and a child. An individual lives for 2 periods. In the first one, she receives an education and a bequest. In the second period, she works, has a child, consumes, and provides an education and a bequest for her daughter.

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.\(^3\)

**Assumption.** \( u \in C^2 \) satisfies

\[ u'(c) > 0, \quad u''(c) < 0, \quad \lim_{c \to 0} u'(c) = +\infty, \quad u(0) = 0, \quad \text{and} \]

\[ u'(c^*) = 1 \quad \text{for some} \; c^* \in \mathbb{R}_+. \]

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 \( c \)). We normalize the interest rate paid on bequests to 1. Education is transformed in future wages through the household production technology \( y(\theta, c; E) \), where \( \theta \in [\theta_0, \theta_1] \) is each child’s productivity parameter, \( c \) is the amount of education and \( E \) is the general level of education.\(^4\)

---


\(^3\)The dependence of the parent’s utility function on the child’s wealth rather than on her utility is a 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.

\(^4\)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

$$y_e (\theta, c; E) > 0, \quad y_\theta (\theta, c; E) > 0, \quad y_{c\theta} (\theta, c; E) > 0, \quad y_E (\theta, c; E) > 0,$$

$$y (\theta, 0; E) > 0, \quad y_{ce} (\theta, c; E) < 0, \quad \lim_{e \to 0} y_e (\theta, c; E) = +\infty,$$

$$\lim_{e \to 0} y_e (\theta, c; E) = 0, \quad \lim_{e \to \infty} y_E (\theta, c; E) < k.$$

The assumption $y_E (\theta, c; E) > 0$ means that education is a source of a positive externalities.$^5$ This is the most interesting case in the model presented although the results trivially hold when there is no externality in education (i.e. $y_E (\theta, c; E) = 0$). The assumption $y_{c\theta} (\theta, c; E) > 0$ means that education increases earnings more for abler individuals. $y_e > 0$ and $y_{ce} < 0$ mean that education increases earnings in a decreasing fashion, while $y (\theta, 0; E) > 0$ means that someone with no education is still able to earn some positive salary.$^6$

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

$$x = y (\theta, c; E) + t. \quad (1)$$

The mother’s wealth, denoted by $Y$, is itself a function of her education which is predetermined in the period we study. Let $\Gamma$ be the space of possible wealth levels. We define $h (Y, \hat{c})$ as the probability function of $Y$ given the educational profile of the previous generation $\hat{c}$. As the parent’s education is predetermined, we omit the term $\hat{c}$ 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

$$Y = c + ke + t. \quad (2)$$

Let $\phi (\theta) \in C^0$ be the probability function of $\theta$. The following assumption ensures that the government is unable to rule out some realizations of $\theta$.

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

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

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

$^5$ See Blaug (1965, pp. 234-241), Cohn (1979) or Lucas (1988) for discussions on the presence of human capital externalities.

$^6$ The assumption $\lim_{e \to \infty} y_E < k$ 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.
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, non-diversifiable, and hard to collateralize implying that private credit markets may fail to finance education. In this economy, credit markets are imperfect in the sense that individuals cannot borrow to finance education.\(^7\)

The household’s problem is

\[
\max_{(e,t)} u(Y - t - ke) + \gamma(\theta, e; E) + t \quad \text{s.t. } t \geq 0.
\]

As usual, the parent will choose the amount of education such that its marginal cost \(ku'(c)\) equals its marginal benefit \(y_e(\theta, c; E)\). 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.\(^8\)

Solving the household’s problem we obtain the following proposition:

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

\[
e(\theta, Y) = \max\{e^u(\theta, Y, k, E); e^c(\theta, Y, k, E)\},
\]

\[
t(\theta, Y) = \max\{Y - e^c - ke(\theta, Y, k, E); 0\},
\]

\[
c(\theta, Y) = \min\{e^c; Y - ke(\theta, Y, k, E)\}.
\]

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

\[
ku'(Y - t - ke) = y_e(\theta, e; E),
\]

\[
u'(Y - t - ke) = 1 + \mu,
\]

\[
\min\{t, \mu\} = 0.
\]

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

\[
y_e(\theta, e^u; E) = k,
\]

\[
Y - e^c - ke^u = t^u.
\]

\(^7\)This is a usual assumption in education and child labour models. See, for example, Baland and Robinson (2000) or Ranjan (2001).

\(^8\)The existence of \(e^u\) and \(e^c\) is demonstrated in the Appendix.
If $\mu > 0$, we say that the solution to the problem is constrained and it follows that:

$$ y_e (\theta, e^*; E) = k u' (Y - ke^*) > k \cdot e^* < e^u, $$
$$ u' (Y - ke^*) > 1 \cdot e^* < e^u. $$

**Remark 2** As can be seen in the proof above, if $Y < e^* + ke^u$, the household’s decisions are constrained since the parent would prefer to leave negative bequests but she is not allowed to. Then, she partially reduces her consumption and partially reduces her daughter’s education. Since education is increasing in $\theta$, households with sufficiently bright children (high $\theta$) or low wealth $Y$ are constrained.⑨

### 4 The Government Intervention Solution

#### 4.1 The First-Best Solution

As in most public finance literature, we take a government that maximizes the unweighted sum of every individual’s utilities. The total level of education is defined as the sum of each individual’s education:⑩

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

Then, the first-best solution (or, equivalently, the symmetric Pareto optimal allocation) is the solution to the following problem:

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

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

⑨Applying the implicit function theorem, we get:

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

⑩This specification implies in 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 in the PhD.). However, as would be clear when we present the decentralized scheme, the main results do not depend on such assumption.
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 on 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 $c^*$ implicitly by the relation

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

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

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

The assumption above guarantees that there are enough resources so that $c^*$ and $e^*$ are feasible in a symmetric information economy.\(^{11}\)

**Proposition 3** The first-best allocations are

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

**Proof.** The first order conditions (necessary and sufficient) for the problem above are

$$[-u' (Y - t (\theta, Y) - kc (\theta, Y)) + 1] h (Y) \phi (\theta) = 0,$$

$$[-k u' (Y - t (\theta, Y) - kc (\theta, Y)) + y_e (\theta, c (\theta, Y), E) + \lambda] h (Y) \phi (\theta) = 0,$$

$$\bar{E} [y_E (\theta, c; E)] - \lambda = 0,$$

where $\lambda$ is the Lagrange multiplier associated to (3). Solving these equations, we get the result above. \(\blacksquare\)

**Remark 4** 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 in the sense that individuals with the same ability receive the same education.\(^{12}\) 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.

Moreover, positive externalities imply that an inefficiently low amount of education is provided in the laissez-faire competitive equilibrium even for unconstrained households (since $y$ is strictly concave in $c$).

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).\(^{13}\)

\(^{11}\)The existence of $e^*$ is demonstrated in the Appendix.

\(^{12}\)We also get that the optimal consumption level $c^*$ is independent of $Y$.

\(^{13}\)Applying the implicit function theorem, we get $\frac{\partial e^* (\theta, Y)}{\partial Y} = - \frac{\partial y_e}{\partial c^*} > 0$. 

7
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.\(^{14}\) We assume that \( Y \) is observable but \( \theta \) 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 utility function, it can be written as

\[
U (\theta, Y) = u (Y - \tau (Y) - f (\theta, Y)) + y (\theta, e (\theta, Y); E) - m (\theta, Y). \quad (5)
\]

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 a local condition and a monotonicity condition.

**Lemma 5** 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

\[
U_\theta (\theta, Y) = y_\theta (\theta, e (\theta, Y); E), \quad (6)
\]

\[
e_\theta (\theta, Y) \geq 0, \quad (7)
\]

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. Define \( P (\theta, Y, E) \) as the utility obtained in the laissez-faire equilibrium. Then, the household’s utility must satisfy

\[
U (\theta, Y) \geq P (\theta, Y - \tau (Y), E), \forall Y, \forall \theta. \quad (8)
\]

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 each period there is a generation of young (paying \( \tau (Y) + f (\theta, Y) \) as taxes and receiving \( ke (\theta, Y) \) in education) and old individuals (paying the deferred

\(^{14}\)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 these type of contracts.
payments $m(\theta, Y)$ and we allow the government to transfer resources between them. Hence, the government budget constraint is

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

Equation (9) states that the net tax revenue is enough to finance the educational expenses.

The government problem is:

$$\max_{\{c^*, \tau, f, m, E\}} \int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) \left[ u(Y - \tau(Y) - f(\theta, Y)) + g(\theta, e(\theta, Y); E) - m(\theta, Y) \right] \phi(\theta) dY d\theta$$

s.t. (3), (5), (6), (7), (8), (9).

Solving the government problem, we get the following proposition, whose proof is presented in the Appendix.

**Proposition 6** The optimal educational policy implements the first-best amount of education and consumption $\{c^*(\theta, Y), c^*; \theta \in [\theta_0, \theta_1], Y \in \Gamma\}$ and achieves first-best welfare.

The basic intuition behind this result is that when 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 while the indirect utility of an individual who purchases education privately decreases. Hence, the participation constraint can be costlessly implemented. Moreover, since the individuals are risk neutral, any redistribution of wealth does not change the total welfare.

**Remark 7** As shown in proposition 6, when transfers between generations are allowed, the optimal educational policy provides equality of opportunities in education (since $c^*(\theta, Y)$ 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 Pigouian taxes and public provision of credit. This implementation is desirable due to its simplicity and informational advantage.
4.3 The 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. 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.

Let $\tau(Y), f(\theta,Y)$ and $m(\theta,Y)$ be the taxes defined before. Define $t(\theta,Y), f$ and $\hat{k}$ as

$$t(\theta,Y) = -m(\theta,Y),$$

$$f(\theta,Y) = t(\theta,Y) + \hat{k}e(\theta,Y),$$

$$\tau(Y) = f.$$ 

Hence, we are restricting the mechanism offered to a lump sum tax $f$, a linear in education up-front fee $t(\theta,Y) + \hat{k}e(\theta,Y)$ and a deferred payment $-m(\theta,Y)$ (which is a subset of the class of contracts considered previously). This mechanism can be alternatively interpreted as a loan $-t(\theta,Y)$ and an up-front fee $\hat{k}e(\theta,Y)$. Clearly, allowing for loans makes it possible to relax the non-negative bequests constraint.

In general, $\hat{k}$ could depend on $\theta$ and $Y$ and $f$ could depend on $Y$. However, as we show below, they are both constant under the optimal policy.

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

$$\int_{\theta_{0}}^{\theta_{1}} \int_{\Gamma} h(Y) \left[ \left(\hat{k} - \hat{k}\right) e(\theta,Y) - f \right] \phi(\theta)dYd\theta \leq 0. \quad (10)$$

Substituting these instruments in the household’s budget constraint, it follows that the total amount of consumption, bequests, and taxes must be equal to the household’s wealth:

$$Y = c + t(\theta,Y) + f + \hat{k}e(\theta,Y). \quad (11)$$

Hence, we can write the household’s problem as:

$$\max_{\left\{e(\theta,Y),t(\theta,Y)\right\}} u \left(Y - t(\theta,Y) - f - \hat{k}e(\theta,Y)\right) + y(\theta,e(\theta,Y);E) + t(\theta,Y).$$

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

\footnote{See Baunol (1972) and Kopeck (2003).}
daughter. Hence, each parent must be consuming $c^*$. Moreover, the marginal benefit of education $y_c$ must be equal to its marginal cost $k$. This result is stated formally in the following lemma:

**Lemma 8** The solution to the household’s problem is \( \{ c^P(\theta, Y), e^P(\theta, Y), \)
\( t^P(\theta, Y); \theta \in [\theta_0, \theta_1], Y \in \Gamma \} \), such that:
\[
c^P(\theta, Y) = c^*,
\]
\[
\hat{k} = y_c(\theta, e^P(\theta, Y); E),
\]
\[
t^P(\theta, Y) = Y - c^* - f - \hat{k}e^P(\theta, Y).
\]

**Proof.** The result follows from the first order conditions of the household’s problem.

Now, we are ready to show that the first-best solution can be reached through a suitable choice of $f$ 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 $E[y_E]$. The lump-sum tax is set in order to cover the expenses from the subsidies. Therefore, it must be equal to the average subsidy $E[e] E[y_E].$

**Proposition 9** The first-best welfare is reached in the asymmetric information context through an appropriate choice of an up-front fee. Moreover, this equilibrium satisfies the government’s budget constraint (10).

**Proof.** Set $\hat{k}$ as
\[
\hat{k} = k - E[y_E(\theta, c^*, E(e^*))].
\]
Substituting in the first order conditions of the household’s problem (13), we get $e(\theta, Y) = e^*(\theta, Y)$.

Set $f$ as
\[
f = E[e^*] E[y_E(\theta, c^*, E(e^*))].
\]
Then, it follows that
\[
E[t] = E[Y - \hat{k}e^* - c^* - f] = E[Y - ke^* - c^*] = E[t^*].
\]
Hence, as $c^*$, $e^*(\theta, Y)$, and $E[t^*]$ are the same as in the first-best solution (and utility is linear in $t$), first-best welfare is achieved.

---

\[16\] The existence of $e^P$ is demonstrated in the Appendix.
From (14), it follows that
\[
\left( k - \hat{k} \right) e^* (\theta, Y) = E \left[ y_E (\theta, e^*, \hat{E} (e^*)) \right] e^* (\theta, Y).
\]

Applying $\hat{E}$ to both sides of the above expression yields
\[
\hat{E} \left[ \left( k - \hat{k} \right) e^* \right] = \hat{E} \left[ y_E (\theta, e^*, \hat{E} (e^*)) \right] \hat{E} [e^*]. \tag{16}
\]

Hence, (15) and (16) imply that $\hat{E} \left[ \left( k - \hat{k} \right) e^* - f \right] = 0$. It follows that the government’s budget constraint (10) is satisfied. ■

**Remark 10** As $E_{t} [t] = \hat{E}_{t} [t^*] > 0$, the government transfers resources from older individuals to younger individual (who repay when older).

Define the household’s financial contribution as
\[
z(\theta, Y) \equiv f + \hat{k} e (\theta, Y).
\]

As education is independent of wealth, it’s 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 brighter children contribute more than those with brighter children.

### 4.3.1 The implementation when the government has access to a foreign market

By assuming that the government can freely transfer wealth between generations, we take a steady-state analysis (whereby steady-state means constant distribution of taxes and educational benefits).\(^{17}\) However, if we start from an economy that has not adopted this policy yet, then the old generation won’t have agreed on a deferred payment $m (\theta, Y)$ and we should ask the best way to implement this policy.\(^{18}\) If there are enough resources to compensate for the lack of deferred payments in the first period then the answer is trivial. In this section, we show that if the government has access to a foreign credit market, then we also do not have any problems in implementing the optimal policy.

Define $D^t (\theta, Y)$ as the government’s deficit with the child whose ability is $\theta$ and parental income is $Y$ at period $t$ :
\[
D^t (\theta, Y) = \left( k - \hat{k} \right) e (\theta, Y) - f \text{ at time } t.
\]

\(^{17}\)Since the distribution of the ability parameter is stationary, stationarity of education implies in stationarity of that wealth distribution.

As $u$ and $y$ are concave and the government can transfer wealth linearly across generations through taxes the optimal policy must be stationary.

\(^{18}\)This point is similar to the discussions on the transition from pay-as-you-go to fully-funded systems of social-security in overlapping-generations models.
As the interest rate is normalized to 1, the government’s intertemporal budget constraint is
\[ \bar{E} \left[ \sum_{t=0}^{\infty} (D^t - t^t + t^{t+1}) \right] \leq 0. \]

Assume that under the efficient education, the distribution of \( Y \) is already stationary (i.e. assume that the first generation has received the efficient amount of education). Since \( u \) and \( y \) are concave, and the government can transfer wealth linearly through the credit market, it follows that \( D^t (\theta, Y) \) must be constant in \( t \). Hence, a necessary and sufficient condition for the government debt to be sustainable is \( \bar{E} [D] \leq 0 \) which is the budget constraint in the problem solved above.

Thus, when the government has access to a foreign market, the decentralized mechanism derived above achieves first-best and satisfies the government budget constraint. As \( \bar{E} [t] = \bar{E} [t^*] > 0 \), the government borrows from abroad to finance education and consumption. Each individual repays her loans in the following period.

5 The economy with default

In the model presented, the returns to education are deterministic. In reality, however, an individual can neither be sure about finishing his education successfully nor about his future returns after a successful conclusion. Hence, educational returns display a very high variation since students may not graduate or not find a job.\(^{19}\)

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.

The optimal educational policy in this case is similar to the income-contingent policies discussed at Krueger and Bowen (1993) and Barr (1991).

Let \( \psi: [\theta_0, \theta_1] \times \Gamma \rightarrow [0, 1] \) be the proportion of individuals with type \( \theta \) and parent’s wealth \( Y \) who are able to repay the deferred payments \( m(\theta, Y) \geq 0 \).

\(^{19}\)The high variation of educational returns was originally pointed out at Becker (1964, pp. 104). For a recent study on this issue, see Miller and Volker (1993).
We assume that $\psi(\theta, Y) > 0$. Although no restrictions on the dependence of $\psi(\theta, Y)$ on $Y$ are needed for our results, it may seem reasonable to assume that $\psi$ is increasing in $Y$.

The household’s utility function is:

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

(17)

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

Define the function $\tilde{U} : [\theta_0, \theta_1]^2 \times \Gamma \to \mathbb{R}$ as

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

Then, the incentive compatibility constraint is

$$U(\theta, \theta, Y) \geq U(\tilde{\theta}, \theta, Y), \quad \forall \tilde{\theta}, \theta.$$ 

As in Lemma 5, the incentive compatibility constraint can be written as a local condition and a monotonicity condition.

**Lemma 11** 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

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

(18)

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

(19)

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

$$E[\tau(Y) + f(\theta, Y) + \psi(\theta, Y)m(\theta, Y) - t(\theta, Y)] \geq E[ke(\theta, Y)].$$

(20)

Then, the government faces the following problem:

$$\max_{\{c, \tau, f, m, E\}} \int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) U(\theta, Y) \phi(\theta) dY d\theta$$

s.t. (3), (8), (17), (18), (19), (20),

$$m(\theta, Y) \geq 0,$$

$$t(\theta, Y) \geq 0.$$ 

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

**Proposition 12** The optimal educational policy in the economy with default implements the first-best amount of education and consumption $\{c^*(\theta, Y), c^*; \theta \in [\theta_0, \theta_1], Y \in \Gamma\}$ 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 labour income allows her to.

**Remark 13** If $\psi$ is interpreted as the probability that the labour income of a type-$0$ 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.

In this model, we have assumed that the supply of labor does not depend on education (possibly because labor is supplied inelastically). One might argue that when the supply of labor is endogenous, moral hazard could emerge. However, in this case, the incentive problem would be the opposite to the one we have considered since individuals with lower ability would prefer to hide their private information. Then, in this case, the existence of tests and exams would probably mitigate that problem.

6 Conclusion

In this paper, we show that the possibility of adverse selection caused by the unobservability of ability can be fully mitigated when the government is able to transfer wealth between generations. Hence, the results obtained by De Fraja (2002) are a consequence of the assumption that the government’s budget constraint must be balanced with each generation at every period. As efficiency requires equality of marginal productivity of education across individuals and the optimal educational policies are Pareto efficient in this case, it follows that the amount of education does not depend on parental income (i.e. equality of opportunities in education is provided).

The inefficiency of the laissez-faire equilibrium was due to two problems: imperfect credit markets and educational externalities. We have shown that the government should provide credit in order to correct the credit market inefficiency. Governmental provision of credit is probably the most suggested educational policy.20 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 education causes on the rest of the economy, the amount of education each household provides in the laissez-faire equilibrium is inefficiently low. We show that the government may obtain the first-best solution 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.\footnote{Friedman (1955, pp.124-125) advocated for a scheme similar to the Pigouvian taxes proposed here. He argued that since buyers of education generate external benefits on those not purchasing education, the government should subsidize those purchasing education and tax those who are not.}

Hence, the optimal mechanism can be decentralized through Pigouvian taxes and credit provision at the market interest rate. An advantage of decentralization is that it requires less information: the government may not know each household’s wealth and the distributions of ability and wealth (it is sufficient to know the optimal externality level and the social marginal benefit it causes).

Moreover, each household’s financial contribution does not depend on income increases in the ability of the child. Thus, the optimal educational policy is not regressive (i.e., wealthier households do not contribute less than poorer households).\footnote{However, it is still input regressive and output regressive inasmuch as education and utility increase in ability (see Arrow, 1971).}

When the returns to education are random, the non-transferability of human capital implies in the emergence of default. In this case, the optimal educational policy can be implemented by some type of income-contingent loans.

**A Appendix**

**A.1 Proof of Lemma 5:**

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

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

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 $\theta$:

$$
\frac{\partial \hat{U} (\theta, \theta, Y)}{\partial \hat{\theta}} \bigg|_{\hat{\theta} = \theta} = 0
$$

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

The first-order condition yields, for almost all $\theta$,

$$
-w' (Y - \tau (Y) - f (\theta, Y)) f_0 (\theta, Y) + y_c (\theta, e (\theta, Y) ; E) e_\theta (\theta, Y) - m_\theta (\theta, Y) = 0.
$$
Differentiating the first-order condition, we get
\[-u'(Y - \tau(Y) - f(\theta, Y)) f_{\theta\theta}(\theta, Y) + u''(Y - \tau(Y) - f(\theta, Y)) |f_\theta(\theta, Y)|^2 + y_{e\theta}(\theta, e(\theta, Y); E) e_\theta(\theta, Y) + y_{e e}(\theta, e(\theta, Y); E) [e_\theta(\theta, Y)]^2 + y_e(\theta, e(\theta, Y); E) e_{\theta\theta}(\theta, Y) - m_{\theta\theta}(\theta, Y) = 0.\] (21)

The second-order condition yields, for almost all \(\theta\),
\[-u'(Y - \tau(Y) - f(\theta, Y)) f_{\theta\theta}(\theta, Y) + u''(Y - \tau(Y) - f(\theta, Y)) |f_\theta(\theta, Y)|^2 + y_{e e}(\theta, e(\theta, Y); E) [e_\theta(\theta, Y)]^2 + y_e(\theta, e(\theta, Y); E) e_{\theta\theta}(\theta, Y) - m_{\theta\theta}(\theta, Y) \leq 0.\] (22)

Substituting (21) in equation (22), 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 \(e_\theta(\theta, Y) \geq 0\).

Differentiating equation (5), yields
\[U_\theta(\theta, Y) = -u'(Y - \tau(Y) - f(\theta, Y)) f_\theta(\theta, Y) + y_\theta(\theta, e(\theta, Y); E) + y_e(\theta, e(\theta, Y); E) e_\theta(\theta, Y) - m_\theta(\theta, Y).\]

Substituting the first-order condition in this expression, we get
\[U_\theta(\theta, Y) = y_\theta(\theta, e(\theta); E).\]

This proves the necessity of (6) and (7).

To prove the sufficiency of (6) and (7), assume that a type \(\theta\) strictly prefers to announce \(\bar{\theta} \neq \theta\):
\[\bar{U}(\bar{\theta}, \bar{\theta}, Y) > \bar{U}(\bar{\theta}, \theta, Y).\]

This equation can be rewritten as \(\int_{\bar{\theta}}^{\hat{\theta}} \bar{U}_1(x, \theta, Y) \, dx > 0\), where \(\bar{U}_1(x, \theta, Y) \equiv \partial U(\bar{\theta}, \theta, Y) / \partial \theta\). As \(\bar{U}_1(x, x, Y) = 0\) for almost all \(x\), it follows that
\[\int_{\bar{\theta}}^{\hat{\theta}} \left[\bar{U}_1(x, \theta, Y) - \bar{U}_1(x, x, Y)\right] \, dx = \int_{\bar{\theta}}^{\hat{\theta}} \int_x^{\theta} \bar{U}_{12}(x, z, Y) \, dz \, dx > 0.\]

As \(\bar{U}_{12}(\bar{\theta}, \theta, Y) = y_{e\theta}(\theta, e(\bar{\theta}, Y); E) e_\theta(\bar{\theta}, Y) \geq 0\) and \(x\) is between \(\theta\) and \(\bar{\theta}\), this inequality cannot hold. \(\blacksquare\)
A.2 Proof of Proposition 6:

From (9), we can substitute \( m(\theta, Y) \) in the welfare function

\[
W = \int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) \left[ u(Y - \tau(Y) - f(\theta, Y) + y(\theta, e(\theta, Y); E) + \tau(Y) + f(\theta, Y) - ke(\theta, Y) \right] \phi(\theta) dY d\theta.
\]

Substituting (5) in the above expression, it follows that

\[
W = \int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) [U(\theta, Y) + m(\theta, Y) + \tau(Y) + f(\theta, Y) - ke(\theta, Y)] \phi(\theta) dY d\theta.
\]

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

\[
\begin{align*}
\dot{S}(\theta) &= \int_{Y \in \Gamma} h(Y) e(\theta, Y) \phi(\theta) dY, \quad (23) \\
S(\theta_0) &= 0, \quad S(\theta_1) = E.
\end{align*}
\]

For the moment, we will ignore the monotonicity condition (7). Let \( Y \in \Gamma \) be an arbitrary wealth level. Then, the optimal policy offered to an individual with wealth \( Y \) must solve the following Hamiltonian:

\[
H = \int_{Y \in \Gamma} h(Y) [U(\theta, Y) + m(\theta, Y) + \tau(Y) + f(\theta, Y) - ke(\theta, Y)] \phi(\theta) dY \\
+ \rho(\theta) \int_{Y \in \Gamma} h(Y) e(\theta, Y) \phi(\theta) dY + \gamma(\theta, Y) y_{\theta} e(\theta, Y; E) \\
+ \lambda(\theta, Y) [u(Y - \tau(Y) - f(\theta, Y)) + y(\theta, e(\theta, Y; E) - m(\theta, Y) - U(\theta, Y)] \\
+ \mu(\theta, Y) \{U(\theta, Y) - P(\theta, Y - \tau(Y), E)\}.
\]

The control variables are \( m, f, \tau, \) and \( e \) and the state variables are \( U \) and \( S \).
\( \rho(\theta), \gamma(\theta), \lambda_{\theta}(\theta), \) and \( \mu_{\theta}(\theta) \) are the multipliers associated with (23), (6), (5), and (8), respectively.

The first order conditions are

\[
\frac{\partial H}{\partial m(\theta, Y)} = 0 \Rightarrow h(Y) \phi(\theta) = \lambda(\theta, Y), \quad (24)
\]

\[
\frac{\partial H}{\partial f(\theta, Y)} = 0 \Rightarrow Y - \tau(Y) - f(\theta, Y) = c^*, \quad (25)
\]

\[
\frac{\partial H}{\partial \tau(Y)} = 0 \Rightarrow \mu(\theta, Y) P_{Y}(\theta, Y - \tau(Y), E) = 0 \Rightarrow \mu(\theta, Y) = 0, \quad (26)
\]

\[
\frac{\partial H}{\partial e(\theta, Y)} = 0 \Rightarrow -kh(Y) \phi(\theta) + \rho(\theta) h(Y) \phi(\theta) + \gamma(\theta, Y) y_{\theta} e(\theta, e(\theta, Y; E) + \lambda(\theta) y_{e} e(\theta, e(\theta, Y; E) = 0, \quad (27)
\]

18
\[
\frac{\partial H}{\partial U(\theta, Y)} = -\gamma_\theta (\theta, Y) : \gamma (\theta, Y) = \gamma (Y) \text{ constant in } \theta,
\]  
(28)

\[
\frac{\partial H}{\partial S(\theta)} = -\rho (\theta) : \rho (\theta) = \rho \text{ constant for all } \theta,
\]  
(29)

\[
\min\{\mu (\theta, Y); U (\theta, Y) - P (\theta, Y - \tau (Y), E)\} = 0.
\]  
(30)

From equation (25), it follows that the first-best amount of consumption is provided in the relaxed problem.

**Lemma 14** \(\gamma (Y) = 0\) for almost all \(Y\).

**Proof.** As \(U (\theta_1, Y)\) is free for all \(Y\), the transversality condition is \(\gamma (\theta_1, Y) = 0\). Hence, (28) implies in \(\gamma (Y) = 0\) for almost all \(Y\).

Substituting \(\gamma (Y) = 0\) in equation (27), yields

\[
y_c (\theta, e (\theta, Y); E) = k - \rho.
\]  
(31)

**Lemma 15** 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 6. As \(\frac{\partial W}{\partial E}|_{E=E^{2b}} = \rho\), it follows that

\[
\int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) y_E (\theta, e_i^{2b} (\theta), E^{2b}) \phi (\theta) dY d\theta = \rho.
\]

Substituting in (31), we get

\[
y_c (\theta, e (\theta, Y); E) = k - E [y_E (\theta, e^{2b}, E^{2b})],
\]

which is the equation that implicitly defines \(e^* (\theta)\).

Hence, the amount of education and consumption solving the relaxed problem is the same as in the first-best solution. Since individuals utilities are linear in deferred payments \(m (\theta, Y)\), it follows that any profile of deferred payments such that the government’s budget constraint is satisfied as an equality achieves the same welfare \(W\). Therefore, first-best welfare is reached in the relaxed problem.

Notice that \(P (\theta, 0, E) = 0\) and \(P_Y (\theta, Y - \tau (Y), E) \geq 1\). 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 (30) is satisfied.

We have to show that the monotonicity condition (7) is satisfied in the relaxed problem considered. But, as we have already shown, \(e^* (\theta)\) is increasing in \(\theta\). Therefore the monotonicity condition \(e_\theta (\theta, Y) \geq 0\) is satisfied.
A.3 Proof of Proposition 12:

Define $S(\theta)$ as in equation (23). Analogously, we introduce the auxiliary variables $x(\theta, Y) = m_\theta(\theta, Y)$, $z(\theta, Y) = c_\theta(\theta, Y)$. Hence, the government’s problem can be written as

$$\begin{align*}
\max_{\theta_0} & \int_{\theta_0} \int_{Y \in \Gamma} h(Y) \left[ u(Y - \tau - f) + y + \tau + f - kc \right] \phi(\theta) dY d\theta \\
\text{s.t.} & \quad \dot{S}(\theta) = \int_{Y \in \Gamma} h(Y) \psi_\theta(\theta) dY, \\
& \quad U_\theta = y_0 - \psi_\theta m, \\
& \quad m_\theta = x, \\
& \quad c_\theta = z, \\
& \quad U = u(Y - \tau - f) + y - \psi m + t, \\
& \quad y_\theta z - \psi_\theta x \geq 0, \\
& \quad U \geq P(\theta, Y - \tau, E), \\
& \quad m \geq 0, \\
& \quad t \geq 0,
\end{align*}$$

where the state variables are $U, c, m, S$, the control variables are $f, \tau, x, t, z$ and we omit the dependence on $\theta, Y$ and $E$ for notational clarity. Setting up the Hamiltonian, we get

$$H = \int_{Y \in \Gamma} h(Y) \left[ u(Y - \tau - f) + y + \tau + f - kc \right] \phi(\theta) dY + \mu^1(\theta) \int_{Y \in \Gamma} h(Y) \psi_\theta(\theta) dY + \mu^2(\theta, Y) [y_0 - \psi_\theta m] + \mu^3(\theta, Y) x + \mu^4(\theta, Y) z$$

$$- \mu^5(\theta, Y) [u(Y - \tau - f) + y - \psi m + t - U] + \lambda^1(\theta, Y) [y_\theta z - \psi_\theta x]$$

$$+ \lambda^2(\theta, Y) [U - P(\theta, Y - \tau, E)] + \lambda^3(\theta, Y) m + \lambda^4(\theta, Y) t(\theta, Y).$$

From the first order conditions, it follows that:

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

$$\lambda^2(\theta, Y) = 0, \text{ for almost all } Y, \theta, \quad (32)$$

$$\mu^3(\theta, Y) = \lambda^1(\theta, Y) \psi_\theta(\theta, Y), \quad (33)$$

$$\rho^1(\theta, Y) = \lambda^4(\theta, Y), \quad (34)$$

$$\mu^4(\theta, Y) = -\lambda^1(\theta, Y) y_\theta(\theta, e(\theta, Y); E), \quad (35)$$

$$\rho^1(\theta, Y) = -\mu^2(\theta, Y), \quad (36)$$

$$- \mu^5(\theta, Y) = h(Y) \phi(\theta) \left[ y_\theta(\theta, e(\theta, Y); E) - k + \mu^1(\theta, Y) \right]$$

$$+ \mu^2(\theta, Y) y_\theta(\theta, e(\theta, Y); E) - \rho^1(\theta, Y) y_\theta(\theta, e(\theta, Y); E)$$

$$+ \lambda^1(\theta, Y) y_\theta(\theta, e(\theta, Y); E) z(\theta, Y), \quad (37)$$

$$\mu^1(\theta) = \mu^1 \text{ constant,} \quad (38)$$

20
\[-\mu_3^2 (\theta, Y) = -\mu^2 (\theta, Y) \psi_\theta (\theta, Y) + \rho^1 (\theta, Y) \psi (\theta, Y) + \lambda^3 (\theta, Y). \quad (40)\]

Substituting (32) in (37),

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

Notice that equation (33) 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

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

Then, as \(\lambda^4 (\theta, Y) \geq 0\), from equation (35) we get \(u' (Y - \tau (Y) - f (\theta, Y)) \geq 1\), for almost all \(Y, \theta\).

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

\(Y - \tau (Y) - f (\theta, Y) = c^e\), 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 (37),

\[\mu^2 (\theta_1, Y) = \rho^1 (\theta, Y) = 0, \text{ for almost all } Y, \theta. \quad (41)\]

Then, from equation (40), we get

\[\mu_3^3 (\theta, Y) = -\lambda^3 (\theta, Y) \leq 0. \quad (42)\]

As \(m (\theta_0, Y)\) and \(m (\theta_1, Y)\) are free, the transversality conditions impose that \(\mu_3^3 (\theta_0) = \mu_3^3 (\theta_1) = 0\). Hence,

\[\mu^3 (\theta, Y) = \lambda^3 (\theta, Y) = 0, \text{ for almost all } Y, \theta. \quad (43)\]

Therefore, equations (34) and (3) imply \(\lambda^4 (\theta, Y) = \mu^4 (\theta, Y) = 0\). Then, from equation (38) we get

\[\mu^1 = k - y_e (\theta, e (\theta, Y); E). \quad (44)\]

Let \(e^{2b}, E^{2b}\) be the amounts of education and externalities that solve the problem above. Then, as in Lemma 15,

\[\frac{\partial W}{\partial E}^{e=e^{2b}, E=E^{2b}} = \mu^1 (\theta_1) = \mu^1.\]

Therefore, it follows that

\[\int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h (Y) y_E (\theta, e^{2b} (\theta, Y), E^{2b}) \phi (\theta) dY d\theta = \mu^1.\]
Substituting into (44),
\[ y_e (\theta, e^{2b} (\theta, Y); E) = k - \int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) y_E (\theta, e^{2b} (\theta), E^{2b}) \phi (\theta) dY d\theta, \]
which is the equation that defines \( e^* \).

As in proposition 6, it’s always possible to satisfy the participation constraint through an appropriate variation in \( \tau (Y) \) and \( f (\theta, Y) \).

Hence, the amount of education and consumption solving the problem above is the same as in the first-best solution. Since individuals 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. \( \blacksquare \)

### A.4 Existence of the equilibria:

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

**Proposition 16** 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 \( \xi : \mathbb{R}_+ \rightarrow \mathbb{R}_+ \) as \( \xi (e) \equiv y_e (\theta, e; E) \). Then, as \( \xi \) is continuous, \( \lim_{e \rightarrow 0} \xi (e) = +\infty \) and \( \lim_{e \rightarrow +\infty} \xi (e) = 0 \), it follows that there exists \( e^u \) such that \( \xi (e^u) = k \). Moreover, as \( \xi'(e) = y_{ce} (\theta, e; E) < 0 \), \( e^u \) is unique.

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

The same argument establishes the existence of the education profile defined before Lemma 8.

**Corollary 17** There exists a unique \( e^p \) such that \( \hat{k} = y_e (\theta, e^p (\theta, Y); E) \).

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

**Proposition 18** There exists a unique \( e^* \) such that
\[ k = y_e (\theta, e^* (\theta, Y); E [e^*]) + \hat{E} [y_e (\theta, e^*, E (e^*))]. \]

**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 \( \{ e (\hat{\theta}) ; \hat{\theta} \neq \theta \} \).

Define the function \( \rho \) as
\[ \rho (e (\theta), e (-\theta), E) \equiv y_e (\theta, e (\theta); E) + \hat{E} [y_e (\theta, e (-\theta), E)] - k. \]
Then, as $\lim_{\epsilon \to 0} \rho(e, e(-\theta), E) = +\infty$, $\lim_{\epsilon \to +\infty} \rho(e, e(-\theta), E) < 0$ and $\rho$ 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 $\lim_{\epsilon \to +\infty} \rho(e, e(-\theta), E) < 0$ and $\rho$ is continuous, there exists $\tilde{e}$ such that, for all $e > \tilde{e}$, $\rho(e, e(-\theta), E) < 0$.

Define $\epsilon$ and $P$ as $\epsilon \equiv [0, \tilde{e}]$ and $P \equiv \{e(\theta) \in \Theta \cap [\theta_0, \theta_1] \}$. 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 $T(E, e) = (\tilde{E}, \tilde{e})$, 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 $\tilde{e}$, it follows that $\tilde{E} \in P$).

Then, the Schauder-Tychonoff Theorem implies the existence of a fixed point of $T$, $(\tilde{E}, \tilde{e})$ (see, Dunford and Schwartz, 1988, p. 456). From the definition of $T$, this fixed point must satisfy equation (4) and $\tilde{E} = \int_{\theta_0}^{\theta_1} \tilde{e}(\theta) \phi(\theta) d\theta$.

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

Acknowledgements.

We wish to thank Luis Henrique Braido, Daniel Ferreira, Andrew Horowitz, and Rodrigo Soares for helpful comments. Remaining errors are all ours.

References


[34] Psacharopoulos, G., 1986. Financing Education in Developing Countries. World Bank, Washington D.C.


25
<table>
<thead>
<tr>
<th>Número</th>
<th>Título</th>
<th>Autor(es)</th>
<th>Data</th>
<th>Páginas</th>
</tr>
</thead>
<tbody>
<tr>
<td>462.</td>
<td>POLÍTICA DE COTAS E INCLUSÃO TRABALHISTA DAS PESSOAS COM DEFICIÊNCIA</td>
<td>Marcelo Côrtes Neri; Alexandre Pinto de Carvalho; Hessia Guilhermo Costilla</td>
<td>Novembro de 2002</td>
<td>67 págs.</td>
</tr>
<tr>
<td>465.</td>
<td>ATIVOS E SAÚDE NO BRASIL</td>
<td>Marcelo Côrtes Neri; Wagner L. Soares</td>
<td>Dezembro de 2002</td>
<td>29 págs.</td>
</tr>
<tr>
<td>466.</td>
<td>INFLAÇÃO E FLEXIBILIDADE SALARIAL</td>
<td>Marcelo Côrtes Neri; Maurício Pinheiro</td>
<td>Dezembro de 2002</td>
<td>16 págs.</td>
</tr>
<tr>
<td>467.</td>
<td>DISTRIBUTIVE EFFECTS OF BRAZILIAN STRUCTURAL REFORMS</td>
<td>Marcelo Côrtes Neri; José Márcio Camargo</td>
<td>Dezembro de 2002</td>
<td>38 págs.</td>
</tr>
<tr>
<td>468.</td>
<td>EMPLOYMENT AND PRODUCTIVITY IN BRAZIL IN THE NINETIES</td>
<td>José Márcio Camargo; Marcelo Côrtes Neri; Maurício Cortez Reis</td>
<td>Dezembro de 2002</td>
<td>32 págs.</td>
</tr>
<tr>
<td>470.</td>
<td>CUSTO DE CICLO ECONÔMICO NO BRASIL EM UM MODELO COM RESTRIÇÃO A CRÉDITO</td>
<td>Bárbara Vasconcelos Boavista da Cunha; Pedro Cavalcanti Ferreira</td>
<td>Janeiro de 2003</td>
<td>21 págs.</td>
</tr>
<tr>
<td>471.</td>
<td>THE LONG-RUN ECONOMIC IMPACT OF AIDS</td>
<td>Pedro Cavalcanti Ferreira; Samuel de Abreu Pessoa</td>
<td>Janeiro de 2003</td>
<td>31 págs.</td>
</tr>
<tr>
<td>472.</td>
<td>A GENERALIZATION OF JUDD’S METHOD OF OUT-STeady-STATE COMPARISONS IN</td>
<td>Paulo Barelli; Samuel de Abreu Pessoa</td>
<td>Fevereiro de 2003</td>
<td>7 págs.</td>
</tr>
<tr>
<td>474.</td>
<td>THE LONG-RUN ECONOMIC IMPACT OF AIDS</td>
<td>Pedro Cavalcanti G. Ferreira; Samuel de Abreu Pessoa</td>
<td>Fevereiro de 2003</td>
<td>30 págs.</td>
</tr>
</tbody>
</table>

A NOTE ON COLE AND STOCKMAN - Paulo Barelli; Samuel de Abreu Pessoa – Abril de 2003 – 8 págs.


<table>
<thead>
<tr>
<th>Código</th>
<th>Título</th>
<th>Autor(es)</th>
<th>Data</th>
<th>Páginas</th>
</tr>
</thead>
<tbody>
<tr>
<td>494</td>
<td>ELASTICITY OF SUBSTITUTION BETWEEN CAPITAL AND LABOR: A PANEL DATA APPROACH</td>
<td>Samuel de Abreu Pessoa; Silvia Matos Pessoa; Rafael Rob</td>
<td>Agosto de 2003</td>
<td>30 págs</td>
</tr>
<tr>
<td>495</td>
<td>A EXPERIÊNCIA DE CRESCIMENTO DAS ECONOMIAS DE MERCADO NOS ÚLTIMOS 40 ANOS</td>
<td>Samuel de Abreu Pessoa</td>
<td>Agosto de 2003</td>
<td>22 págs</td>
</tr>
<tr>
<td>496</td>
<td>NORMALITY UNDER UNCERTAINTY</td>
<td>Carlos Eugênio E. da Costa</td>
<td>Setembro de 2003</td>
<td>08 págs</td>
</tr>
<tr>
<td>497</td>
<td>RISK SHARING AND THE HOUSEHOLD COLLECTIVE MODEL</td>
<td>Carlos Eugênio E. da Costa</td>
<td>Setembro de 2003</td>
<td>15 págs</td>
</tr>
<tr>
<td>498</td>
<td>REDISTRIBUTION WITH UNOBSERVED 'EX-ANTE' CHOICES</td>
<td>Carlos Eugênio E. da Costa</td>
<td>Setembro de 2003</td>
<td>30 págs</td>
</tr>
<tr>
<td>499</td>
<td>OPTIMAL TAXATION WITH GRADUAL LEARNING OF TYPES</td>
<td>Carlos Eugênio E. da Costa</td>
<td>Setembro de 2003</td>
<td>26 págs</td>
</tr>
<tr>
<td>500</td>
<td>AVALIANDO PESQUISADORES E DEPARTAMENTOS DE ECONOMIA NO BRASIL A PARTIR DE CITAÇÕES INTERNACIONAIS</td>
<td>João Victor Issler; Rachel Couto Ferreira</td>
<td>Setembro de 2003</td>
<td>29 págs</td>
</tr>
<tr>
<td>501</td>
<td>A FAMILY OF AUTOREGRESSIVE CONDITIONAL DURATION MODELS</td>
<td>Marcelo Fernandes; Joachim Grammig</td>
<td>Setembro de 2003</td>
<td>37 págs</td>
</tr>
<tr>
<td>502</td>
<td>NONPARAMETRIC SPECIFICATION TESTS FOR CONDITIONAL DURATION MODELS</td>
<td>Marcelo Fernandes; Joachim Grammig</td>
<td>Setembro de 2003</td>
<td>42 págs</td>
</tr>
<tr>
<td>503</td>
<td>A NOTE ON CHAMBERS’S “LONG MEMORY AND AGGREGATION IN MACROECONOMIC TIME SERIES”</td>
<td>Leonardo Rocha Souza</td>
<td>Setembro de 2003</td>
<td>11 págs</td>
</tr>
<tr>
<td>504</td>
<td>ON CHOICE OF TECHNIQUE IN THE ROBINSON-SLOW-SRINIVASAN MODEL</td>
<td>M. Ali Khan</td>
<td>Setembro de 2003</td>
<td>34 págs</td>
</tr>
<tr>
<td>505</td>
<td>ENDOGENOUS TIME-DEPENDENT RULES AND THE COSTS OF DISINFLATION WITH IMPERFECT CREDIBILITY</td>
<td>Marco Bonomo; Carlos Viana de Carvalho</td>
<td>Outubro de 2003</td>
<td>27 págs</td>
</tr>
<tr>
<td>507</td>
<td>TESTING PRODUCTION FUNCTIONS USED IN EMPIRICAL GROWTH STUDIES</td>
<td>Pedro Cavalcanti Ferreira; João Victor Issler; Samuel de Abreu Pessoa</td>
<td>Outubro de 2003</td>
<td>8 págs</td>
</tr>
<tr>
<td>508</td>
<td>SHOULD EDUCATIONAL POLICIES BE REGRESSIVE?</td>
<td>Daniel Gottlieb; Humberto Moreira</td>
<td>Outubro de 2003</td>
<td>25 págs</td>
</tr>
<tr>
<td>509</td>
<td>TRADE AND CO-OPERATION IN THE EU-MERCOSUL FREE TRADE AGREEMENT</td>
<td>Renato G. Flôres Jr.</td>
<td>Outubro de 2003</td>
<td>33 págs</td>
</tr>
</tbody>
</table>