Should Educational Policies be Regressive?*

Daniel Gottlieb,† Humberto Moreira‡

July 19, 2010

Abstract

In this paper, we study optimal educational policies when the ability to benefit from education is private information. We extend the framework of De Fraja (2002) in two directions. First, we replace his unusual specification of the government’s budget constraint, which prevents the government to use tax revenues from an older generation to subsidize the education of a younger generation, by the usual one. We show that the optimal educational policies achieve the first-best, are not regressive, and can be decentralized through Pigouvian taxes and credit provision. Second, we consider utility functions that are not quasi-linear. In this case, we show that the first-best can no longer be reached, education may not be monotonic in ability and progressivities of education are locally welfare improving.

Keywords: Education, Redistribution, Optimal Taxation.
JEL Classification: I22, H21, H23, H52.

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.¹ Economists have usually argued that reducing the existing regressivity of education would result in significant welfare gains.²

The trade-off between efficiency and equity plays a fundamental role in the analysis of optimal tax policies. In the specific case of education, this issue was originally addressed

*We wish to thank Luis Henrique Braido, Carlos E. da Costa, Daniel Ferreira, Andrew Horowitz, Rodrigo Soares, Thierry Verdier, and specially Antoine Bommier, Pierre Dubois, and James Poterba for helpful comments and suggestions. We would also like to thank seminar participants at Getulio Vargas Foundation, UFRJ, Toulouse, the 2004 North American Summer Meeting of the Econometric Society, the 2004 Meeting of the European Economic Association, and the 2003 Meeting of the Brazilian Econometric Society. As usual, all remaining errors are ours.

¹Department of Economics, MIT.
²Graduate 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.
3See, for example, Fernandez and Rogerson (1996), Kozol (1991) or Psacharopoulous (1986).
4Fernandez and Rogerson (1998) argue that the welfare gains from the equalization of spending across districts in American states are about 3.2 percent of steady-state income.
by Hare and Ulph (1979), who introduced schooling in Mirrlees’ (1971) original model and assumed that the ability to benefit from it was only partially observable by the government. Fleurbaey et al. (2002) extended the Hare and Ulph model to the case where the ability to benefit from education is unobservable and the government has a Rawlsian welfare function.

In a recent contribution, De Fraja (2002), hereafter DF, studied the optimal educational provision in an overlapping-generations model in the presence of externalities and imperfect capital markets. His surprising results suggest that educational policies: (i) should be regressive (in the sense that households with brighter children and higher incomes contribute less in absolute terms than those with less bright children and lower incomes), (ii) should 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), and (iii) should be input-regressive (meaning that education should be increasing in ability). Therefore, the regressivity of educational systems in most countries may actually reflect the optimality of such policies, and attempts to provide an equality of opportunities in education may generate large efficiency losses.

In this paper, we show that the results obtained by DF critically rely on two specific assumptions which are hard to justify and are not common in the taxation literature: (a) a nonstandard government budget constraint and (b) quasi-linear utility functions combined with a utilitarian welfare function.

The regressivity of the financing mechanism and the inequality of opportunities (results i and ii mentioned above) are caused by a nonstandard restriction on the government’s budget constraint: It is assumed that deferred payments made by the older generation cannot be used to finance the education of the younger generation. However, the cross-subsidization of education is one important feature of public educational systems. Moreover, this restriction on the budget constraint generates a welfare loss. Hence, there seems to be neither positive nor normative justifications for this assumption.

When we allow the government to finance education using deferred payments from the older generation in the DF framework, 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 implemented in a decentralized way through a competitive equilibrium with 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 tax 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 can also be implemented in an environment where household’s wealth is unobservable.

Unlike argued by DF, the reason for the regressivity and the inequality of opportunities of the second-best policy is not the trade-off between efficiency and equity but one between efficiency and rent-extraction. Since the utility function is quasi-linear and the government

---

3 More specifically, they assumed that although the government could not condition financial contributions on the ability to benefit from education, it could use this information for allocational purposes.

4 According to DF: “The regressivity of the optimal education system derived in the paper can be interpreted as implying that pursuing redistributive goals using education policies is bound to have a substantial cost in terms of the sub-optimality of the education policy implemented: the fundamental message of the paper is that there is a stark conflict between equity and efficiency in education.”
has a utilitarianist welfare function, there is no preference for redistribution and the government simply minimizes deadweight loss.\textsuperscript{5} The need to use education in order to extract rents arises simply due to the unusual restriction that prevents the use of revenues from deferred payments to finance education. Then, since poorer individuals are those who benefit most from education, minimizing deadweight loss implies that the education of poorer individuals should be taxed more.\textsuperscript{6}

Since the utility function is assumed to be linear in the wealth left to the next generation, the optimal policy generates a large inequality of the future generation’s wealth and no inequality of consumption. However, it seems counter-intuitive that a society who cares about the redistribution in the current period would not care at all about inequality of the wealth of the future generation.

We show that when the utility functions are strictly concave in the wealth of the future generation, input-regressive policies may no longer be incentive-compatible. Moreover, the trade-off between inefficiency and equity prevents the first-best welfare from being achievable.

The non-implementability of input-regressive policies result is fairly general as it does not depend on the specific welfare function considered. Hence, it applies not only to the DF model, but also to Rawlsian welfare functions as considered by Fleurbaey et al. (2002). Therefore, our model suggests that, unlike in the models of DF and Fleurbaey et al., it may be the case that education cannot be increasing in ability and an attempt to implement a monotonic educational policy may instead lead to welfare losses.\textsuperscript{7}

The remainder of the paper is organized as follows. Section 2 presents the general framework. Subsection 2.1 discusses the government’s budget constraint and 2.2 defines the competitive equilibrium without government intervention. Subsections 2.3 and 2.4 define the government-intervention solution when the ability to benefit from education is and is not observable, respectively. Section 3 characterizes the solutions when the utility function is quasi-linear and Section 4 considers the case where the utility function is strictly concave. Then, Section 5 summarizes the main results of the paper.

## 2 General Framework

We consider an overlapping generations model where individuals live for 2 periods. In the first period, the individual 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 child. Each household consists of a parent and a child. There is a continuum of households with measure normalized to 1 in each period. As we will focus on steady-state equilibria, time subscripts will be omitted.

Each individual’s utility function is

\[ U = u(c) + v(x), \]

\textsuperscript{5}Provided that the aggregate income is such that it is socially optimal for someone to leave positive bequests.  
\textsuperscript{6}The assumption of positive education externalities implies that the uniform commodity taxation result does not apply and it is optimal to tax education and consumption at different rates even though the utility function is separable.  
\textsuperscript{7}The technical reason for the possible non-implementability of input-regressive policies is that the single-crossing condition may not hold when the utility function in strictly concave. Both Fleurbaey et al. and DF assume quasi-linear utility functions, which ensures that the single-crossing condition is satisfied.
where $c$ is her consumption, $x$ is the amount of monetary resources available to the child, and $u$ and $v$ are strictly increasing, twice continuously differentiable functions, and $\lim_{c \to 0} u'(c) = +\infty$.\textsuperscript{8} Section 3 considers the case where $v$ is linear whereas Section 4 considers the case where $v$ is strictly concave.

There are two ways of transferring wealth to a child: bequests $t$ and higher future wages (through education $e$). The rate of return of bequests is normalized to 1. Education is transformed in future wages through a production technology $y(\theta, e; E)$, where $e$ is the amount of education, $\theta \in [\theta_0, \theta_1]$ is each child’s ability to benefit from education, and $E$ is the general level of education.\textsuperscript{9} We assume that $\theta$ is continuously distributed on $[\theta_0, \theta_1]$ with a strictly positive density $\phi(\theta)$.

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

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

The mother’s wealth, denoted by $Y$, is itself a function the amount of education she received in her childhood. Let $\Gamma \subset \mathbb{R}_+$ be the space of possible wealth levels. Let $h(Y, \tilde{e})$ denote the probability function of $Y$ given the educational profile of the previous generation $\tilde{e}$. We consider the optimal educational policies for the current generation taking as fixed the amount of education of the previous generation. Therefore, 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. The household’s budget constraint is

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

We assume that the production technology $y : [\theta_0, \theta_1] \times \mathbb{R}_+^2 \to \mathbb{R}_+$ is twice continuously differentiable and satisfies the Inada conditions $\lim_{e \to 0} y_e(\theta, e; E) = +\infty$ and $\lim_{e \to +\infty} y_e(\theta, e; E) = 0$.\textsuperscript{10} Furthermore, it satisfies the following conditions:

$$y_e(\theta, e; E) > 0 \quad (A1), \quad y_{ee}(\theta, e; E) < 0 \quad (A2), \quad y_{\theta}(\theta, 0; E) > 0 \quad (A3),$$

$$y_{e\theta}(\theta, e; E) > 0 \quad (A4), \quad y_E(\theta, e; E) \geq 0 \quad (A5), \quad \lim_{e \to +\infty} y_E(\theta, e; E) < k. \quad (A6).$$

Assumptions (A1) and (A2) state that education increases wages in a decreasing rate. Assumption (A3) states that wages are increasing in the ability to benefit from education $\theta$. Assumption (A4) means that the marginal returns to education increases in the individual ability. Assumption (A5) states that education may be a source of positive externalities. Assumption (A6), which states that the externalities are not large enough to exceed the cost of education as the amount of education approaches infinity, is useful to ensure the existence of a solution.

\textsuperscript{8}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.

\textsuperscript{9}Notice that in the model presented, it is immaterial if education serves only as a screening device or whether it enhances productivity.

\textsuperscript{10}The Inada conditions are helpful for establishing the existence of equilibria.
For notational convenience, define the expectations operator $\bar{E} [\cdot]$ as

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

The general level of education is defined as the sum of every individual’s education:\footnote{This specification implies that any unit of education generates the same externality (i.e., the externality of a year in high school is the same as a year preparing for the PhD). However, it is straightforward to generalize our results to the case where $E = \bar{E} [g(e)]$, for a strictly increasing function $g : \mathbb{R}_+ \to \mathbb{R}_+$.}

$$
E = \bar{E} [e]. \tag{3}
$$

2.1 The Government’s Budget Constraint

Following DF, we assume that the government 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 student loans. The parent’s wealth $Y$ is observable but the daughter’s ability to benefit from education $\theta$ is private information.

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

$$
Y = c(\theta, Y) + \tau(Y) + f(\theta, Y). \tag{4}
$$

In each period a parent with wealth $Y$ and whose child has ability $\theta$ 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. In DF, it is imposed that the government is unable to use revenues from deferred payments to finance education. Hence, the government budget constraint is

$$
\bar{E} [\tau + f] \geq \bar{E} [ke], \quad \bar{E} [m] \geq 0,
$$

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

As deferred payments and education are the only channels of transferring wealth between generations (since bequests have been normalized to zero), imposing that the aggregate amount of deferred payments be non-negative is equivalent to restricting all transfers between generations to be done through the provision of education. However, contributions for public education account for a substantial amount of intergenerational transfers. For example, Bommier et al. (2004) have shown that when expenditures in public education are included in generational accounts along with pensions and health care, generations usually thought to have benefitted from intergenerational transfers were actually net losers.

The only justification provided for such specification is given in footnote 11:
"Of course future taxes could also be used to repay the debt, but this would reduce the amount available for the education of future generations, and must therefore be subject to a limit, normalized here to 0; this value is consistent with a steady state analysis, where the taxes levied on the future generation must correspond to those levied on the current generation to pay for the previous generation’s education."

This argument does not explain why the government would not be able to use revenues from deferred payments to finance education (or consumption) since, in each period, all taxes are paid by the same generation (namely, by the parents). Hence the assumption that the budget constraint is balanced with each generation in each period is also consistent with a steady state analysis. Moreover, since the latter is the usual assumption in overlapping-generation models, is weaker than DF’s specification, and is consistent with the evidence on the importance of intergenerational transfers for public education, it does not seem reasonable to further restrict the government’s problem.\footnote{Furthermore, the choice of 0 for the limit of deferred payments allowed to be spent in education is not just a normalization. As will be clear in Section 2.4, the first-best is obtained even under the unusual budget constraint specified by DF if this limit is lower than the total cost of first-best expenditures on education minus the first-best amount of income taxes.}

We will allow the government to finance education through deferred payments. Therefore, the government’s budget constraint is

\begin{equation}
\bar{E} [\tau + f + m] \geq \bar{E} [ke].
\end{equation}

Equation (5) states that the net tax revenue is enough to finance the educational expenditures in each period.

Subsections 2.2, 2.3, and 2.4 define the equilibrium without government intervention, the first-best problem and the second-best problem, respectively.

### 2.2 The Equilibrium without Government Intervention

It is usually argued that investment in human capital is risky, nondiversifiable, 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 leave negative bequests: \( t \geq 0 \).

The parent’s program is

\begin{equation}
\max_{c,e,t} u(c) + v (y (\theta, c; E) + t)
\end{equation}

subject to \( c = Y - ke - t, \quad t \geq 0. \)

A competitive equilibrium without government intervention is a profile of consumption, education, and bequests \( \{c (\theta, Y), e (\theta, Y), t (\theta, Y); \quad \theta \in [\theta_0, \theta_1], \quad Y \in \Gamma\} \) such that: (i) \( c (\theta, Y), e (\theta, Y), t (\theta, Y) \) solve Program (6) for all \( \theta, Y, \) and (ii) \( E = \bar{E} [e] \).
2.3 The First-Best Problem

As in most public-finance literature, we take a utilitarian government. Moreover, we assume that the utility of future generations enter the welfare function only through the weight attributed to it by the current generation.

As usual, we will refer to the outcome chosen by the government if ability were observable as a first-best solution. A first-best solution is a profile of consumption, education, and bequests \( \{c(\theta,Y),e(\theta,Y),t(\theta,Y); \theta \in [\theta_0,\theta_1], Y \in \Gamma\} \) that maximizes the sum of each parent’s utilities subject to the resource constraint and the definition of the general level of education:

\[
\max_{\{c(\theta,Y),e(\theta,Y),t(\theta,Y); \theta \in [\theta_0,\theta_1], Y \in \Gamma\}} \mathbb{E} [u(c) + v(y(\theta,c;E) + t)]
\]

subject to \( \mathbb{E}[Y - c - ke - t] \geq 0 \) and (3).

2.4 The Second-Best Problem

When the government cannot observe each individual’s ability to benefit from education, it has to design an educational system that induces people to self-select. From the revelation principle, the search for an optimal educational system can be restricted to the class of mechanisms that satisfy the following incentive-compatibility constraint:

\[
\theta \in \arg \max_{\hat{\theta} \in [\theta_0,\theta_1]} u(Y - \tau(Y) - f(\hat{\theta},Y)) + v(y(\theta)e(\hat{\theta},Y);E) - m(\theta,Y), \tag{7}
\]

for all \( \theta, \hat{\theta}, Y \).

Following DF, we assume that individuals are not forbidden to purchase education in the private sector. Hence, they will only join the education program when their utility exceeds the utility obtained if they purchase education privately:

\[
u(Y - \tau(Y) - f(\theta,Y)) + v(y(\theta,e(\theta,Y);E) - m(\theta,Y)) \geq P(\theta,Y - \tau(Y),E), \tag{8}
\]

for all \( \theta,Y \), where

\[
P(\theta,Y,E) \equiv \max_{c,t \geq 0} u(Y - t - ke) + v(y(\theta,e;E) + t).
\]

A second-best solution is a profile of education and taxes \( \{\tau(Y), f(\theta,Y), m(\theta,Y), e(\theta,Y); \theta \in [\theta_0,\theta_1], Y \in \Gamma\} \) solving

\[
\max_{\{e,\tau,f,m,E\}} \mathbb{E} [u(Y - \tau - f) + v(y(\theta,e;E) - m)]
\]

subject to (3), (5), (7), and (8).

In Sections 3 and 4, we analyze the cases of quasi-linear and concave utilities separately.

3 Quasi-linear Utility

This Section considers essentially the same economy studied at DF. The only difference is the government budget constraint (5), which allows the government to finance education through
deferred payments. However, as we will show in Subsections 3.2 and 3.3, the optimal policy is drastically different under our budget constraint. The first-best can be implemented with a simple mechanism that does not require the government to observe ability or income.

We assume that the utility function is quasi-linear:

\[ v(x) = x \quad \text{and} \quad u''(c) < 0. \]

Furthermore, we assume that

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

Note that \( c^* \) is the amount of consumption whose marginal utility is equal to the marginal utility of leaving resources to one’s child.

### 3.1 The First-Best Solution

The social marginal benefit of education is the sum of the private marginal benefit of education \( y \varepsilon \) and the social benefit of education \( \bar{E}[y \varepsilon] \). Hence, the first-best amount of education should equate the marginal benefit of education to its marginal cost \( k \). Thus, the first-best profile of education is implicitly determined by the equation:\(^{13}\)

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

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, the first-best amount of education is independent of the parent’s wealth \( Y \) (we shall denote \( e^*(\theta, Y) \) as \( e^*(\theta) \) in order to emphasize that it does not depend on \( Y \)). Moreover, since \( c^* \) is independent of \( Y \), the optimal level of consumption is also independent of wealth.

Furthermore, as long as there are enough resources so that \( e^* \) and \( c^* \) are feasible, any profile of bequests yields the same welfare (since the utility function is linear in bequests and the welfare function is utilitarianist). Thus, an optimal profile of bequests consists of the amount of resources left to the household after the expenditures in education and consumption:

\[
t^*(\theta, Y) = Y - ke^*(\theta) - c^*. \tag{9}
\]

As can be seen from the equation above, the marginal propensity to bequeath under this policy is equal to one.

We will assume that resources are sufficiently high so that \( e^* \) and \( c^* \) are feasible. If this assumption is not satisfied, it would be socially optimal for all individuals to leave zero bequests.

**Assumption 1** \( \bar{E}[t^*] \geq 0. \)

**Proposition 1** The first-best allocations are

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

\(^{13}\)The proofs of existence and uniqueness of the solutions obtained in this paper are presented in the appendix.
Proof. See Appendix.

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).\textsuperscript{14}

Remark 1 When \( y_E > 0 \), even parents who are not credit constrained provide an inefficiently low amount of education in the equilibrium without government intervention due to the presence of positive externalities.

3.2 The Second-Best Solution

Substituting the household’s budget constraint (4) in the utility function, it can be written as

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

The following lemma, whose proof can be found at DF, allows us to substitute the incentive-compatibility constraint for the local first- and second-order conditions of the incentive-compatibility program:

**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

\[
\frac{\partial U(\theta,Y)}{\partial \theta} = y_\theta(\theta,e(\theta,Y);E),
\]

\[
\frac{\partial e(\theta,Y)}{\partial \theta} \geq 0,
\]

for all \( \theta,Y \).

The proposition below states 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 education system. Moreover, since this implementation does not require any additional resources, it achieves first-best welfare.

**Proposition 2** 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 the first-best welfare. Furthermore, the optimal profile of deferred payments is such that \( m^*(\theta,Y) = -t^*(\theta,Y) \), where \( t^* \) is as defined in equation (9).

**Proof.** See Appendix.

The basic intuition behind this result is that when the government raises the income tax uniformly and decreases the up-front fee 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

\textsuperscript{14}Applying the implicit function theorem, we obtain \( \frac{\partial e^*(\theta)}{\partial \theta} = -\frac{y_ee}{y_ee} > 0. \)
constraint (8) can be implemented at no cost. Moreover, since the utility function is quasi-linear and the social welfare function is utilitarianist, any redistribution of wealth does not change welfare.\footnote{In that sense, the model is similar to the regulation model when the shadow-cost of public funds is zero. However, unlike in the standard regulation models [e.g. Laffont and Tirole (1993)], the shadow-cost of public funds is endogenous in this model. In DF, it is only zero in the extreme case where the government is able to circumvent the restriction that deferred payments cannot be used to finance education and income taxes (Proposition 6). As Proposition 2 shows, the first best is implementable even when the extreme assumptions of DF are not satisfied if deferred payments can be used to finance education.}

Proposition 2 implies that when we consider budget constraint (5), a strictly higher welfare is achieved (since first-best is implemented). Moreover, contrary to the results of DF, the optimal policy provides equality of opportunities in education (since $e^* (\theta)$ does not depend on $Y$).\footnote{Notice also that the second-best education profile is strictly greater than the one obtained in a competitive equilibrium (see Remark 1).}

Although Proposition 2 states that the first-best can be achieved, it does not show how. In the next subsection, we address the implementation of the optimal policy. It is shown that the first-best can be obtained through Pigouvian taxes and public provision of credit. Moreover, it does not require that the government observes wealth so that our results also hold in an environment where $Y$ is unobservable.

### 3.3 Implementation through Pigouvian taxes

In this subsection, we will restrict the space of contracts to those consisting of lump-sum taxes, a linear up-front 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

\[
t (\theta, Y) = -m (\theta, Y),
\]

\[
\hat{k} (\theta, Y) = \frac{f (\theta, Y) + m (\theta, Y)}{e (\theta, Y)}.
\]

In general, $\hat{k}$ and $t$ could depend on $\theta$ and $Y$. However, as we show below, they are both constant for all $\theta$ and $Y$ under the optimal policy. Hence, an optimal contract consists of a lump-sum tax $\tau$, a linear 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 $\tau$, 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 (5), it can be written as

\[
\hat{E} \left[ (k - \hat{k}) e - \tau \right] \leq 0. \quad (13)
\]

In each period, the government pays $(k - \hat{k})$ as a subsidy on each unit of education and receives $\tau$ as a lump-sum tax. The government also loans $\hat{E} [\tau]$ 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. In a steady-state, repayments are equal to loans in each period. Thus, as usual, the budget constraint simply states that the total expenses should not exceed the total revenues of the government.

Substituting the definitions of \( \hat{k} \) and \( t \) in the parent’s budget constraint (4), it follows that the total amount of consumption, loans repaid, and taxes must be equal to the household’s wealth:

\[
Y = c(\theta, Y) + t(\theta, Y) + \tau + \hat{k} e(\theta, Y). \tag{14}
\]

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

\[
\max_{e(\theta, Y), t(\theta, Y)} u\left(Y - t(\theta, Y) - \tau - \hat{k} e(\theta, Y)\right) + y(\theta, c(\theta, Y); E) + t(\theta, Y)
\]

subject to \( E = \bar{E}[e] \).

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 child. Hence, each parent must be consuming \( c^* \). Moreover, the private marginal benefit of education \( y_e \) must be equal to its marginal cost \( \hat{k} \). Therefore, we get the following lemma:

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

\[
e^P(\theta, Y) = c^*;
\]

\[
\hat{k} = y_e(\theta, e^P(\theta, Y); E), \tag{15}
\]

\[
t^P(\theta, Y) = Y - c^* - \tau - \hat{k} e^P(\theta, Y).
\]

**Proof.** The result follows from the first-order conditions (which are necessary and sufficient).

Now, we are ready to show that the first-best welfare can be reached through a suitable choice of \( \tau \) and \( \hat{k} \). 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 \( \hat{k} \) 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 \( \bar{E}[e^P] \bar{E}[y_E] \). Substituting into the parent’s problem, we get the following result:

**Proposition 3** There exists a second-best solution where the price of education \( \hat{k} \) and the tax \( \tau \) are both constant in \( \theta \) and \( Y \). Moreover, this solution achieves the first-best welfare.

**Proof.** See Appendix.

The fact that the optimal policy can be implemented with a lump-sum tax \( \tau \), a constant price of education \( \hat{k} \) and a constant interest rate implies that it can also be implemented when household’s wealth is unobservable.

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, \( z(\theta, Y) \) is strictly increasing in ability since \( \hat{k} > 0 \) and \( \hat{e}^*(\theta) > 0 \). Therefore, households with brighter children contribute more than households with less bright children. These results differ from DF, 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.

Let \( x^P \) denote the wealth left to the child:

\[
x^P(\theta, Y) \equiv y(\theta, e^*(\theta); E) + t^P(\theta, Y).
\]

Then, it follows that the marginal propensity to bequeath under the Pigouvian scheme is equal to one (i.e., \( \frac{\partial x^P}{\partial Y} = 1 \)). In other words, every additional amount of wealth is left to the future generation. Furthermore, individuals with higher ability receive more wealth through education than ones with less ability (since \( \frac{\partial x^P}{\partial \theta} = y_\theta > 0 \)).

Therefore, the optimal educational policy generates large inequalities of wealth left to the future generation. The quasi-linearity of the utility function implies that welfare is not affected by this inequality. In the next section, we study how the results change when the utility function is not quasi-linear so that individuals care not only about redistribution of present consumption but also about redistribution of wealth left for future generations.

### 4 Strictly Concave Utility

In the previous Section, we have followed DF in assuming that parents’ preferences can be represented by a utility function linear in the wealth left to their children. This assumption implies that a utilitarian government does not have preferences for redistribution of bequests. Then, as was shown in Subsection 3.3, the optimal policy implements first-best welfare but generates large inequalities of wealth left to the future generation.

In this section, we assume that the parents’ utility function is strictly concave in their children’s wealth:

\[
v''(x) < 0, \quad u''(x) < 0.
\]

This assumption implies that parents are risk-averse in the wealth of their children and a utilitarian government has preference for redistribution. We also assume that \( v \) satisfies the Inada condition \( \lim_{x \to 0^+} v'(x) = +\infty \).

Substituting the household’s budget constraint (4) in the utility function, we obtain

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

Then, it follows that

\[
U_{\theta e} = v'(y - m)[y_{e\theta} - r_A(v, y - m)y_e y_\theta],
\]

where \( r_A(v, x) \equiv -\frac{v''(x)}{v'(x)} \) is the absolute coefficient of risk-aversion and measures the concavity of \( v \).

The sign of \( U_{\theta e} \) is ambiguous since an increasing profile of education has two opposite effects in the parent’s utility. The first effect \( (y_{e\theta} > 0) \) concerns efficiency: an increasing profile of education benefits more those with higher marginal productivity of education. The
second effect ($-r_A y_\theta y_\theta < 0$) concerns equity: an increasing profile of education gives more wealth to those with lower marginal utility. Then, the sign of $U_{e\theta}$ will depend on the preference for redistribution relative to efficiency (captured by the risk-aversion coefficient $r_A$).\textsuperscript{17}

The following lemma presents the standard necessary (second-order) condition for incentive-compatibility.

\textbf{Lemma 3} An incentive-compatible $C^2$ by parts policy \{$\tau(Y), f(\theta, Y), m(\theta, Y), e(\theta, Y); \theta \in [\theta_0, \theta_1], Y \in \Gamma$\} satisfies

\[ U_{e\theta} \frac{\partial e(\theta, Y)}{\partial \theta} \geq 0, \]

for all $\theta, Y$.

\textbf{Proof.} Omitted. \blacksquare

Thus, incentive-compatible policies are generally not input-regressive. When the preference for redistribution is sufficiently small, incentive-compatible policies are input-regressive ($\frac{\partial e(\theta, Y)}{\partial \theta} > 0$) as in the quasi-linear environment. If the concern for redistribution is sufficiently high, input-regressive policies are no longer incentive-compatible. Indeed, incentive-compatible educational policies may also be non-monotonic.

For example, suppose that parents have constant risk-aversion with respect to wealth left to their children $r_A > 0$. Then, any incentive-compatible education profile must be increasing if $\frac{y_\theta}{y_{e\theta}} \geq r_A$ and decreasing if $\frac{y_\theta}{y_{e\theta}} \leq r_A$. Figure 1 shows incentive-compatible profiles when $\frac{d}{d\theta} \left( \frac{y_\theta}{y_{e\theta}} \right) > 0$ and $\frac{d}{d\theta} \left( \frac{y_\theta}{y_{e\theta}} \right) < 0$, respectively.\textsuperscript{18}

\textbf{Figure 1: Incentive-compatible profiles of education}

In any of those cases, however, a government that does not observe ability cannot implement the first-best. The strict concavity of the utility function introduces a trade-off between

\textsuperscript{17}The preference for redistribution of bequests is absent in DF since $r_A$ is zero when the utility function is quasi-linear.

\textsuperscript{18}When $U_{e\theta}$ does not have a constant sign, the single-crossing property does not hold. Araujo and Moreira (2004) and Araujo, Gottlieb, and Moreira (2008) present a method for solving screening models where the single-crossing property does not hold.
efficiency and equity. The policy from Section 3, which implements the efficient amount of education but generates a large inequality of bequests is no longer desirable.

The trade-off between efficiency and equity implies that optimal policies would typically be distorted towards more progressive policies as in standard optimal taxation models [e.g. Mirrlees (1971)]. Indeed, starting from an undistorted economy (i.e. an economy with Pigou-vian taxes and credit provision), a local increase in the progressivity of education is always welfare increasing. This follows from the well-known fact that, in a previously undistorted economy, distortions cause second-order losses and first-order gains in welfare (when the government cares about redistribution).

**Proposition 4** When $v$ is strictly concave, the first-best amount of welfare cannot be reached when the government does not observe the ability to benefit from education $\theta$. Moreover, starting from an undistorted economy, a local increase in the progressivity of education is strictly welfare enhancing.

**Proof.** See Appendix. ■

5 Conclusion

In this paper, we have shown that the regressivity results obtained by DF are driven by the assumptions that the government is not allowed to use deferred payments in education, and that a utilitarian government maximizes quasi-linear utility functions (which implies that it minimizes deadweight loss). Then, the government’s problem is to minimize the inefficiency caused by this nonstandard restriction on its budget constraint. When we allow the government to use deferred payments to finance education, the optimal educational policy achieves the same amount of welfare that could be reached if ability were observable.

By not internalizing the effects that education causes on the rest of the economy, the amount of education that each household provides in the equilibrium without government intervention is inefficiently low. We show that the first-best solution can be implemented through Pigouvian taxes and credit provision when neither ability nor wealth are observable. In this context, the appropriate Pigouvian taxes are educational subsidies that induce households to internalize for the (positive) externalities caused by education.

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, the optimal policy provides equality of opportunities in education. 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.

When utility functions are strictly concave so that individuals care not only about re-distribution of present consumption but also about redistribution of wealth left for future generations, input-regressive policies may not be incentive-compatible. The trade-off between efficiency and equity implies that the first-best welfare can no longer be achieved. Moreover, starting from an undistorted economy, increases in the progressivity of education are locally welfare increasing.

14
Appendix

Proofs

Proof of Proposition 1: Introducing the auxiliary variable \( S(\theta) \), equation (3) can be rewritten as

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

Substituting the resource constraint in the objective function, we can write the following Hamiltonian:

\[
H = [u(c(\theta, Y)) + y(\theta, e(\theta, Y)), E) + Y - c(\theta, Y) - ke(\theta, Y)] \phi(\theta) h(Y)
+ \mu(\theta) \int_{Y \in \Gamma} e(\theta, Y) \phi(\theta) h(Y) dY,
\]

where \( c \) and \( e \) are control variables and \( S \) is a state variable. The first-order conditions are

\[
u'(c(\theta, Y)) = 1, \quad y_e(\theta, e(\theta, Y), E) - k = \mu(\theta), \quad \text{and} \quad \mu(\theta) = \mu \text{ constant}.
\]

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

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

Substituting in the first-order conditions, concludes to proof. \( \blacksquare \)

Proof of Proposition 2: Substituting the government’s budget constraint (5) into the objective function, we obtain:

\[
W = \bar{E}[u(Y - \tau - f) + y(\theta, e; E, \tau + f - ke)]
\]

Using equation (10), we can rewrite the objective function as:

\[
W = \bar{E}[U - m + \tau + f - ke]
\]

As in the proof of Proposition 1, rewrite equation (3) using the auxiliary variable \( S(\theta) \) as:

\[
\dot{S}(\theta) = \int_{Y \in \Gamma} e(\theta, Y) \phi(\theta) h(Y) dY, \quad S(\theta_0) = 0, \quad S(\theta_1) = E.
\]  \( (16) \)

For the moment, we will ignore the monotonicity condition (12). Latter, we will verify that it is satisfied in the solution of this relaxed program. Then, we can set up the following Hamiltonian:

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

The control variables are \( m, f, \tau, \) and \( e \) and the state variables are \( U \) and \( S \). \( \rho(\theta), \gamma(\theta, Y), \lambda(\theta, Y), \) and \( \mu(\theta, Y) \) are the multipliers associated with (16), (11), (10), and (8), respectively.

The first-order conditions are:\(^{19}\)

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

\(^{19}\)We omit the dependence of \( y \) on \( \theta, e(\theta, Y), \) and \( E \) for notational clarity.
\[ \frac{\partial H}{\partial f(\theta,Y)} = 0 : Y - \tau(Y) - f(\theta,Y) = c^*, \quad (17) \]

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

\[ \frac{\partial H}{\partial e(\theta,Y)} = 0 : [\rho(\theta) - k] h(Y) \phi(\theta) + \gamma(\theta,Y) y_{e} + \lambda(\theta,Y) y_{e} = 0, \quad (18) \]

\[ \frac{\partial H}{\partial U(\theta,Y)} = -\frac{\partial U(\theta,Y)}{\partial \theta} : \gamma(\theta,Y) \text{ constant in } \theta, \quad (19) \]

\[ \frac{\partial H}{\partial S(\theta)} = -\dot{\rho}(\theta) : \rho(\theta) = \rho \text{ constant in } \theta, \text{ and} \]

\[ 0 = \min \{ \mu(\theta,Y); U(\theta,Y) - P(\theta,Y - \tau(Y),E) \}. \quad (20) \]

From equation (17), it follows that the consumption that solves the program above is equal to the first-best level of consumption \( c^* \). Since \( U(\theta_1,Y) \) is free for all \( Y \), the transversality condition for \( \gamma(\theta,Y) \) is \( \gamma(\theta_1,Y) = 0 \). Therefore, equation (19) yields

\[ \gamma(\theta,Y) = 0, \text{ for almost all } \theta,Y. \]

Substituting into equation (18), we obtain

\[ y_{e}(\theta,e(\theta,Y);E) = k - \rho. \quad (21) \]

Next, we will show that the amount of education solving the relaxed program above is the same as in the first-best solution. In order to do so, let \( e^r \) and \( E^r \) be the amounts of education and externalities that solve the relaxed problem. By the envelope theorem, \( \frac{\partial W}{\partial E} \bigg|_{e=e^r,E=E^r} = \rho \). Therefore,

\[ \int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} y_{E}(\theta,e^r(\theta,Y);E^r) h(Y) \phi(\theta) dY d\theta = \rho. \]

Substituting in equation (21), we obtain

\[ y_{e}(\theta,e^r(\theta,Y);E) + \bar{E}[y_{E}(\theta,e^r,E^r)] = k, \]

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

Note that \( P(\theta,0,E) = 0 \) and \( P_{Y}(\theta,Y,E) \geq 1 \). Moreover, a unitary increase in \( \tau(Y) \) and a unitary decrease in \( f(\theta,Y) \) leaves \( U(\theta,Y) \) unchanged. Hence, it is always possible to choose \( \tau(Y) \) and \( f(\theta,Y) \) such that condition (20) is satisfied.

We have shown that the profiles of education and consumption solving the relaxed problem are the same as the first-best solution. Since the utility functions 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 \( W \). Fix \( \tau(Y) \) and \( f(\theta,Y) \) so that equation (20) is satisfied and let \( m^r(\theta,Y) \) be given by

\[ m^r(\theta,Y) = ke^r(\theta) + c^* - Y. \]

Then, substituting equation (17), we obtain

\[ m^r(\theta,Y) = ke^r(\theta) - \tau(Y) - f(\theta,Y). \]

Taking the expectation of the equation above, we get

\[ \bar{E}[m^r] = \bar{E}[ke^r - \tau - f], \]
which is the government’s budget constraint (equation 5). Therefore, the first-best welfare is reached in the relaxed program.

It remains to be shown that the monotonicity condition (12) is satisfied in the solution of the relaxed program. But, since \( c^*(\theta) \) is increasing in \( \theta \), it follows that this condition is satisfied and, therefore, the relaxed solution also solves the second-best problem.

**Proof of Proposition 3:** Set \( \hat{k} \) as

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

(22)

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

Set \( \tau \) as

\[
\tau = \hat{E} [e^* - \hat{E} [y_E(\theta, e^*, \hat{E}[e^*])]].
\]  

(23)

Then, it follows that

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

Hence, as \( c^*, e^*(\theta) \), and \( \hat{E} \left[ t^* \right] \) are the same as in the first-best solution (and utility is linear in \( t \)), first-best welfare is achieved. From equation (22), we have

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

Applying \( \hat{E} \) to both sides of the above expression, yields

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

(24)

Hence, equations (23) and (24) imply that \( \hat{E} \left[ k - \hat{k} \right] e^* - \tau = 0 \) so that the government’s budget constraint (13) is satisfied.

**Proof of Proposition 4:** Using equation (1), the first-best problem can be written as

\[
\max_{\{c(\theta,Y), e(\theta,Y), x(\theta,Y), \lambda \}} \int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} [u(c(\theta,Y)) + v(x(\theta,Y))] h(Y) \phi(\theta) \, dY \, d\theta
\]

subject to (3), and

\[
\int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} [Y - c(\theta,Y) - ke(\theta,Y) - x(\theta,Y) + y(\theta,e(\theta,Y);E)] h(Y) \phi(\theta) \, dY \, d\theta \geq 0.
\]

Introducing the auxiliary variable \( S(\theta) \), we can write the following Hamiltonian:

\[
H = [u(c(\theta,Y)) + v(x(\theta,Y))] h(Y) \phi(\theta) + \rho(\theta) \int_{Y \in \Gamma} e(\theta,Y) h(Y) \phi(\theta) \, dY
\]

\[
+ \lambda \int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} [Y - c(\theta,Y) - ke(\theta,Y) - x(\theta,Y) + y(\theta,e(\theta,Y);E)] h(Y) \phi(\theta) \, dY \, d\theta.
\]

The control variables are \( c, e, \) and \( x \), whereas the state variable is \( E \). The first order conditions with respect to \( c \) and \( x \) yield

\[
u'(c(\theta,Y)) = v'(x(\theta,Y)) = \lambda.
\]

Therefore, \( c(\theta,Y) \) and \( x(\theta,Y) \) are constant for almost all \( \theta, Y \).
We claim that having \( c(\theta, Y) \) and \( x(\theta, Y) \) constant is not incentive-compatible. Let \( \hat{\theta} > \theta \) and suppose that \( x(\hat{\theta}, Y) = x(\theta, Y) \) and \( c(\hat{\theta}, Y) = c(\theta, Y) \). Then, from equation (1), we must have

\[
y(\hat{\theta}, c(\hat{\theta}, Y); E) + m(\hat{\theta}, Y) = y(\theta, c(\theta, Y); E) + m(\theta, Y).
\]

Since \( \hat{\theta} > \theta \), it follows that

\[
y(\hat{\theta}, c(\hat{\theta}, Y); E) + m(\hat{\theta}, Y) < y(\theta, c(\theta, Y); E) + m(\theta, Y).
\]

If type \( \hat{\theta} \) takes the contract designed for type \( \theta \), she obtains utility:

\[
u(\hat{\theta}, c(\hat{\theta}, Y); E) + v(y(\hat{\theta}, c(\hat{\theta}, Y); E) + m(\hat{\theta}, Y))
\]

where the equality above uses the fact that \( c(\theta, Y) = c(\hat{\theta}, Y) \) and the inequality uses the fact that \( y_\theta > 0 \). Therefore, \( \hat{\theta} \) can benefit from taking the contract designed for type \( \theta \), which violates the incentive-compatibility constraint (7). Hence, the first-best level of welfare cannot be reached when ability is not observable if \( v \) is strictly concave. This concludes the first part of the proof.

As in the proof of Proposition 1, taking that the first order conditions of the Hamiltonian with respect to \( e \) and \( E \) and using the envelope theorem, it follows that the first-best education profile is implicitly determined by:

\[
y_e(c(\theta, Y); E) + E[y_E] = k
\]

In order to establish the second part of the proposition, notice that the derivative of the welfare function with respect to \( c(\theta, Y) \) evaluated at the first-best education profile is

\[
\frac{\partial W}{\partial c(\theta, Y)}|_{c^*, c^*, e^*} = h(Y) \phi(\theta)[v'(x^*(\theta, Y))y_e(c^*(\theta, Y); E) - k],
\]

where \( x^*(\theta, Y) = y(\theta, c^*(\theta, Y); E) + Y - kc^*(\theta) - c^* \). The sign of the derivative above is the same as the sign of \( v'(x^*(\theta, Y))y_e(c^*(\theta, Y); E) - k \), which is decreasing in \( Y \). Recall that, since wealth is observable, there is no incentive-compatibility constraint associated with \( Y \). Therefore, starting from an undistorted education profile, a local increase in progressivity of education increases welfare.

**Existence and Uniqueness of Solutions**

The following proposition ensures that the education profiles in the competitive equilibrium without government intervention (defined in Subsection 2.2) are well defined:

**Proposition 5** Suppose that the utility function satisfies the assumptions of either Section 3 or Section 4. The competitive equilibrium without government exists and is unique.

**Proof.** First, consider the quasi-linear case. It suffices to show that there exist unique \( e^u \) and \( e^c \) such that \( k = y_e(\theta, e^u; E) \), \( ku^e(Y - ke^c) = y_e(\theta, e^c; E) \).

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

Now, consider the strictly concave case. The parent’s program can be written as

$$\max_{t, \hat{c}} u(Y - ke - t) + v(y(\theta, e; \bar{E} + t)).$$

Existence and uniqueness follow immediately from strict concavity and the Inada conditions. ■

The same argument establishes the existence of equilibrium with Pigouvian taxes and credit provision in Subsection 3.3:

**Corollary 1** There exists a unique equilibrium to the decentralized implementation defined in Subsection 3.3.

The following proposition ensures the existence and uniqueness of the first-best level of education in both the quasilinear and strictly concave economies (which is also the second-best level of education in the quasilinear economy).

**Proposition 6** There exists a unique $e^*$ such that

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

(25)

**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

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

Then, as $\lim_{e \to 0} \rho(e(e(-\theta), E) = +\infty$, $\lim_{e \to +\infty} \rho(e(e(-\theta), E) < 0$ (by $\lim_{e \to \infty} y_e(e(\theta); E) = 0$ and (A6)) and $\rho$ is continuous, it follows that, for every $e(-\theta)$ and every $E$, there exists $\hat{e}(\theta)$ such that $\rho(\hat{e}(\theta), e(-\theta), E) = 0$. Moreover, the Inada conditions imply that this $\hat{e}(\theta)$ is unique.

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

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

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

Then, the Schauder-Tychonoff Theorem implies the existence of a fixed point of $T$, $\left( \tilde{E}, \tilde{e} \right)$ (see, Dunford and Schwartz, 1988, p. 456). From the definition of $T$, this fixed point must satisfy equation (25) and $\tilde{E} = \int_{\theta_0}^{\theta_1} \tilde{e}(\theta) \phi(\theta) d\theta$.

Uniqueness follows from the strict concavity of the first-best problem with respect to $e$. ■

---

20The first-best allocations in the quasilinear economy are not unique since the welfare function is indifferent with respect to any distribution of bequests.
References


