Escola de Pós-Graduação em Economia - EPGE
Fundação Getulio Vargas

Two Essays on the Economics of Education

Dissertação submetida à Escola de Pós-Graduação em Economia da Fundação Getulio Vargas como requisito para obtenção do Título de Mestre em Economia

Aluno: Daniel Gottlieb
Orientador: Humberto Luiz Ataíde Moreira

Rio de Janeiro
2004

Escola de Pós-Graduação em Economia - EPGE
Fundação Getulio Vargas

Two Essays on the Economics of Education

Dissertação submetida à Escola de Pós-Graduação em Economia da Fundação Getulio Vargas como requisito para obtenção do Título de Mestre em Economia

Aluno: Daniel Gottlieb

Banca Examinadora:
Humberto Luiz Ataíde Moreira (Orientador, EPGE/FGV)
Aloisio Pessoa de Araújo (EPGE/FGV)
Juliano Junqueira Assunção (PUC/RJ)

Rio de Janeiro
2004

this dissertation is dedicated to Tatiana for her love, support, and encouragement.

Contents

List of Figures

Acknowledgments

The return of an investment in education strongly depends on the quality of teachers and classmates. In that sense, I consider myself extremely lucky for having experienced such a pleasant and intellectually stimulating environment at the Getulio Vargas Foundation over the past two years.

I would like to thank professors Luis Braido, Daniel Ferreira, and Samuel Pessoa with whom I had the opportunity of working as a TA. They, as well as professors Marco Bonomo, Carlos da Costa, and Andrew Horowitz, provided me with countless and fruitful conversations.

I owe an enormous debt of gratitude to my advisor Humberto Moreira and to professor Aloisio Araújo for the opportunity of doing research together.

For comments and suggestions on the first chapter of this dissertation, I wish to thank Luis Braido, Carlos da Costa, Daniel Ferreira, Andrew Horowitz, Marcelo Moreira, Rodrigo Soares, Thierry Verdier, and specially Antoine Bommier, and Pierre Dubois. I would also like to thank seminar participants at IBMEC and the Getulio Vargas Foundation, and the SBE meeting at Porto Seguro, the PAI Network meeting at Toulouse, and the 2004 North American

Summer Meeting of the Econometric Society.

For comments and suggestions on the second chapter, I am thankful to Luis Braido, Carlos da Costa, Flávio Cunha, James Heckman, Derek Neal, Philip Reny, and seminar participants at the University of Chicago and the Getulio Vargas Foundation. I also benefitted from insightful comments from Juliano Assunção.

Finally, I would like to thank my family and, specially, Tatiana, who have supported and encouraged me. They were never tired of discussing my new ideas.

Two Essays on the Economics of Education

Chapter 1

A Model of Mixed Signals with Applications to Countersignaling and the GED

Abstract

We develop a job-market signaling model where signals may convey two pieces of information. This model is employed to study the GED exam and countersignaling (signals non-monotonic in ability). A result of the model is that countersignaling is more expected to occur in jobs that require a combination of skills that differs from the combination used in the schooling process. The model also produces testable implications consistent with evidence on the GED: (i) it signals both high cognitive and low non-cognitive skills and (ii) it does not affect wages. Additionally, it suggests modifications that would make the GED a more effective signal.

1.1 Introduction

Most of the existing signaling models are structured in a way that signals reveal information monotonically. In the job-market models, for example, higher education always discloses information about higher productivity. Nevertheless, in many situations signals convey information about different characteristics. In such cases, good and bad characteristics may be revealed by the same signal so that the monotonicity does not hold (i.e., signals may be mixed).

One example of mixed signals is the General Educational Development (GED) exam, which is taken by high school dropouts to certify their equivalence with high school graduates. The GED reveals, at the same time, high cognitive skills and low non-cognitive skills [Heckman and Rubinstein (2001) and Cavallo, Heckman, and Hsee (1998)]. Moreover, wages received by high school dropouts are not influenced by the realization of this exam.

Another example is the occurrence of countersignaling, where individuals with high types choose to engage in a lower amount of signaling than medium-type individuals. In the context of education as a signal, for example, mediocre individuals appear to educate more than bright individuals for professions where individuals without a licence are not denied work [Hvide (2003)].[1] Unlike standard models of advertising as a signal predict, Clements (2004) documents that many higher-quality products are sold in lower quality packages. In the presence of countersignaling, a higher amount of signal may reveal good or bad information (since high-type and low-type individuals signal less than intermediate types).

A third example of mixed signals is presented by Drazen and Hubrich (2003), where it was argued that higher interest rates show that the government is committed to maintaining the exchange rate fixed, but also signal weak fundaments.

In the initial papers in the signaling literature, the informational asymmetry consisted of a unidimensional parameter which was known only to one side of the market [e.g. Spence (1973, 1974)]. Then, under the natural condition that individuals could be ordered according to their marginal utility of signaling (single-crossing property), there existed a family of separating equilibria, all ranked by the Pareto optimality criterion.[2] Moreover, only the Pareto dominant equilibrium was robust to competition among firms [Riley (1979)].

Of course, the possibility to reduce all asymmetry to a unidimensional parameter is not a very realistic assumption. In the labor market model, for example, this implies that all relevant characteristics of an employee could be captured by a single ability-type, usually thought as a cognitive ability. However, there is significant empirical evidence on the importance of non-cognitive skills as well as cognitive skills in the labor market. Apparently, it was assumed that the generalization of the original results to the multidimensional case would be

[1] A signifficant amount of the 400 richest people in the US do not hold an academic degree (Bill Gates is a well known example) [Orzach et. at (1999)]. Hvide (2003) also argues that many bright MBA students from top-schools dropout to work.

[2] In the specific case of labor market model, this condition implies that education is more costly to less able individuals.

straightforward. This assumption was soon proved wrong by Kholleppel's (1983) example of a two-dimensional extension of Spence's model where no separating equilibrium existed.

Quinzii and Rochet (1985) and Engers (1987) provided sufficient conditions for the existence of a separating equilibrium in the multidimensional model. In Quinzii and Rochet's article, ability was represented by a k-dimensional vector and they assumed the existence of k (non-exclusive) different types of education. Moreover, they assumed that the signaling costs were linear and separable in the signals (up to a change of variables). Hence, it was as if each school required only one type of ability. Then, an individual would be able to attend a school whose system required only a type of skill (cognitive skills, for example) and another school that required only another type of skill (non-cognitive skills). Under this separability assumption (which implies in the single-crossing property in each dimension), Quinzii and Rochet obtain results similar to the unidimensional-characteristic models: only separating equilibria exist and wages are monotonic in the characteristic parameter.

It is needless to say that the educational systems assumed by Quinzii and Rochet are not realistic since all known educational systems require both cognitive and non-cognitive abilities (although in different proportions). Engers relaxed this assumption through a generalization of the unidimensional assumption that individuals' marginal utility of signaling could be ordered (single-crossing property). However, in the multidimensional case, this assumption is much less compelling since, as the number of signals rise, it becomes more probable that the single-crossing property (SCP) does not hold when one controls for one signal (i.e., the introduction of other signals may break the SCP in the multidimensional case).

Hence, the existence of mixed signals contrasts strongly with monotonic wages and separability of types in equilibrium as predicted by standard models. Indeed, when the single-crossing property holds, an equilibrium always exists, signals are always monotonic, and all equilibria are fully-separable.[3] Thus, in order to understand non-monotone signals, the SCP must not be imposed.

In this article, we present a two-dimensional characteristics signaling model where the SCP may not hold. Individuals' characteristics are represented by a vector of cognitive and non-cognitive ability parameters. Firms can access a combination of these characteristics through an interview but cannot precisely tell if the realization of this interview was due to high cognitive or non-cognitive ability. Workers are able to signal their characteristics through the number of years dedicated to education.

This model is employed in order to understand the evidence on the GED and on countersignaling. When applied to the GED, the signaling equilibrium has some interesting properties consistent with the available empirical evidence: individuals with different abilities obtain the same amount of education and passing the exam does not increase one's earnings even though it signals higher

[3] Araujo, Gottlieb, and Moreira (2004a) show that a necessary and sufficient condition for full-separability is that the SCP holds locally.

cognitive skills. These results follow from the fact that GED is a mixed signal: if a worker with low overall ability has passed the exam, it means that his non-cognitive ability is low. Hence, as both types of ability are used in the production process, passing the exam is not necessarily a signal of high productivity.

The model suggests that the problem of the GED exam is its focus on cognitive ability. A test which places a stronger emphasis on non-cognitive ability would be a more effective signal. Moreover, a simple change in the passing standards of the GED would not affect its neutrality on wages.

It is shown that countersignaling occurs whenever the schooling technology differs from the technology of firms. The model has a very intuitive testable implication: the amount of countersignaling is strictly increasing in the difference between the schooling technology and the firms' technology. Hence, countersignaling is expected to be more important in occupations that require a different combination of skills from those required in the schooling process.

The rest of the paper is organized as follows. The basic framework is presented in Section 1.2. Section 1.3 characterizes the equilibrium. Section 1.4 discusses how countersignaling may emerge and Section 1.5 employs this framework to understand the GED exam. Then, Section 1.6 concludes.

1.2 The basic framework

The economy consists of a continuum of informed workers who sell their labor to uninformed firms. Each worker is characterized by a two-dimensional vector of characteristics (ι, η), where ι is her cognitive ability (intelligence) and η is her non-cognitive ability (perseverance). The set of all possible characteristics is the compact set $\Theta \equiv [\iota_0, \iota_1] \times [\eta_0, \eta_1] \subset \mathbb{R}^2_{++}$ and the types are distributed according to a continuous density $p : \Theta \to \mathbb{R}_{++}$, which is assumed to be a C^2 function.

Workers are able to engage in a schooling activity $y \in \mathbb{R}_+$ which firms can observe. By engaging in such activity, the type-(ι, η) worker incurs in a cost $c(\iota, \eta, y)$. Her productivity depends on the vector of innate characteristics which is not (directly) observable.

Firms have identical technologies with constant returns to scale $f(\iota, \eta)$ and act competitively.[4] Moreover, other than schooling, firms have an interview technology $g(\iota, \eta)$ which is a non-sufficient statistic for the worker's productivity. Thus, even though firms have some idea of the overall ability of a worker, they are unable to unambiguously determine her productivity.[5] In a more general model, we could imagine that individuals might exert effort in order to distort the market's assessment of their productivity [e.g. Holmstrom (1999) and Dewatripont, Jewitt, and Tirole (1999)]. This possibility is studied at Araujo,

[4] In this paper, we consider only the pure signaling case since schooling does not influence productivity. However, as was shown in a previous version of the paper, all propositions in this paper also hold when schooling affects productivity [see Araujo, Gottlieb, and Moreira (2004b)].

[5] The hypothesis that firms can access an additional signal which consists of a measure of the worker's ability is also present at Feltovich et al. (2002).

Gottlieb and Moreira (2004b), where it is assumed that schooling distorts the worker's performance in the interview. However, most of the results in this paper still hold.[6]

After observing schooling y and the result of the interview g, each firm offers a wage $w(y,g)$. Thus, each worker will choose the amount of schooling y in order to maximize $w(y,g) - c(\iota,\eta,y)$.

The timing of the signaling game is as follows. First, nature determines each worker's type according to the density function p. Then, workers choose their educational level contingent on their type. Subsequently, firms offer a wage $w(y,g)$ conditional on observing (y,g).

Since all firms are equal, we will study symmetric equilibria where the offered wage schedule is the same for every firm. As usual, we adopt the perfect Bayesian equilibrium concept:

Definition 1 *A perfect Bayesian equilibrium (PBE) for the signaling game is a profile of strategies* $\{y(\iota,\eta), w(y,g)\}$ *and beliefs* $\mu(\cdot \mid y,g)$ *such that*

1. *The worker's strategy is optimal given the equilibrium wage schedule:*

$$(\iota,\eta) \in \arg\max_{\{\tilde{\iota},\tilde{\eta}\}} w\left(y(\tilde{\iota},\tilde{\eta}), g(\tilde{\iota},\tilde{\eta})\right) - c\left(\iota,\eta,y(\tilde{\iota},\tilde{\eta})\right),$$

2. *Firms earn zero profits:* $w\left(y(\iota,\eta), g(\iota,\eta)\right) = E\left[f(\iota,\eta) \mid g,y\right]$.

3. *Beliefs are consistent:* $\mu(\iota,\eta \mid y,g)$ *is derived from the worker's strategy using Bayes' rule where possible.*

Next, we will specify the analytical forms of the functions presented.[7] Consider the following cost of signaling function:

$$c(\iota,\eta,y) = \frac{y}{\iota\eta}. \tag{1}$$

The cost function above implies that intelligence and perseverance are imperfect substitutes in the schooling process.

We assume that the interview function is given by

$$g(\iota,\eta) = \alpha\iota + \eta, \tag{2}$$

where $\alpha > 0$ is the rate of substitution between perseverance and intelligence.

Substituting (2) into (1), we are able to rewrite the cost of signaling as a function of the intelligence and the interview result:

$$c(\iota,g,y) = \frac{y}{\iota(g-\alpha\iota)},$$

where we denote this function by c with some abuse of notation.

[6] It can also be shown that, locally, the ability to distort the result of the interview raises the amount of education in equilibrium for all individuals.

[7] The robustness of the model is studied in the Appendix A.

Notice that, in general, the single-crossing property (SCP) may not be satisfied since

$$c_{y\iota}(\iota, g, y) = -\frac{g - 2\alpha\iota}{\left[\iota(g - \alpha\iota)\right]^2} \left\{ \begin{array}{c} > \\ < \end{array} \right\} 0 \Leftrightarrow \iota \left\{ \begin{array}{c} > \\ < \end{array} \right\} \frac{g}{2\alpha}.$$

The SCP states that the marginal utility of effort is monotonic in the ability parameter. In this specific case, it means that, conditional on the interview g, more intelligence would either always decrease or always increase the cost of schooling.[8] Hence, the SCP is equivalent to assuming that the range of abilities is such that intelligence is always better than perseverance for schooling (or vice-versa).

The intelligence level $\iota = \frac{g}{2\alpha}$ divides the parameter space in two intervals (CS_+ and CS_-) according to the sign of $c_{y\iota}$ (negative and positive, respectively). For workers with intelligence below (above) $\frac{g}{2\alpha}$, intelligence decreases (increases) the cost of signaling given the overall ability g. When the SCP is satisfied, $[\iota_0, \iota_1]$ belongs to one of these intervals.

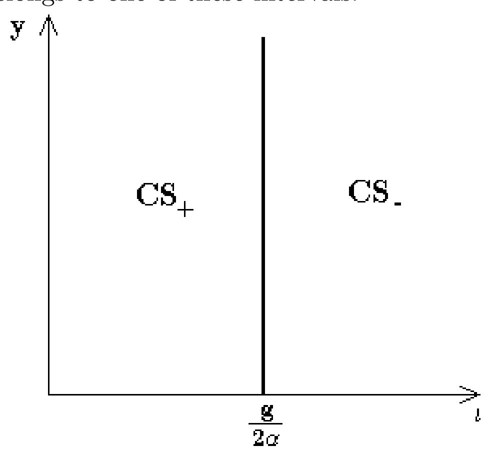


Figure 1

We assume that the worker's productivity is given by the Cobb-Douglas function

$$f(\iota, \eta) = \iota^b \eta^{1-b},$$

where $b \in (0, 1)$. If $b > \frac{1}{2}$ we say that the firm's technology is intensive in cognitive skills. Otherwise, we say that it is intensive in non-cognitive skills.

It is useful to rewrite the production function conditional on the interview g as

$$s(\iota) = \iota^b (g - \alpha\iota)^{1-b}. \tag{3}$$

[8] In other words, even for individuals with very high intelligence and very low perseverance levels, raising a unit of intelligence and decreasing α units of perseverance would decrease the marginal cost of schooling (or the opposite case when the sign of $c_{y\iota}$ is reversed).

1.3 The signaling equilibria

In this section, the signaling equilibrium is characterized. First, we divide the interval of parameters in three different sets according to the degree of separation. Necessary conditions for an equilibrium are presented for each of these sets separately. Then, we present the refinement criterion which will be employed in order to select a unique equilibrium. It consists of a generalization of Riley's (1979) criterion. Subsequently, sufficient conditions for the equilibrium are obtained.

The following definitions will be useful in the characterization of the equilibria.

Definition 2 *Given an equilibrium profile of education y, the pooling set $\Theta(y, g)$ is the set of types whose signal is (y, g).*

We say that a type is separated if, in equilibrium, her characteristics are revealed from her signals y and g. If more than one type choose the same amount of education, we say that they are pooled. As in standard signaling models, an equilibrium may feature a continuum of types choosing the same signal. We call these types continuously pooled. However, when the single crossing property does not hold, the equilibrium may feature non-monotone signalling. As a result, a disconnected set of workers may acquire the same level of education. We say that these workers are discretely pooled. We state the precise definitions below:

Definition 3 *Given an equilibrium profile of education y:*

1. *A type$-(\iota, g)$ worker is separated if $\Theta(y(\iota, g), g) = \{(\iota, g)\}$. A separating set is a set of types where every element is separated.*

2. *A type$-(\iota, g)$ worker is continuously pooled if $\Theta(y(\iota, g), g)$ is not discrete. A continuous pooling set is a set of types where every element is continuously pooled.*

3. *A type$-(\iota, g)$ worker is discretely pooled if $\Theta(y(\iota, g), g) \neq \{(\iota, g)\}$ is discrete. A discrete pooling set is a set of types where every element is discretely pooled.*

In any signaling equilibrium, each type must belong to one of these three sets. In the following subsections, we study the properties of separating sets, continuous pooling sets and discrete pooling sets, respectively.

1.3.1 Separating set

When a worker belongs to a separating set, Bayes' rule implies that belief $\mu(\iota \mid y, g)$ must be a singleton measure concentrated at ι. Then, the zero-profits condition (second condition of Definition 1) is

$$w(y(\iota, g), g) = f(\iota, g - \alpha\iota). \tag{4}$$

The worker's truth-telling condition (first condition) is

$$\iota \in \arg\max_{\{\tilde{\iota}\}} f(\tilde{\iota}, g - \alpha\tilde{\iota}) - c(\iota, g, y(\tilde{\iota}, g)). \tag{5}$$

Notice that each realization of $g(\iota, \eta) = x$ defines a set of possible characteristics

$$g^{-1}(x) \equiv \{(\iota, \eta) \in [\iota_0, \iota_1] \times [\eta_0, \eta_1] : x = \alpha\iota + \eta\}.$$

As the worker's production function is a strictly concave, continuous function of ι, there exists a unique intelligence level such that her productivity is maximal given the overall ability g. This educational level is defined as

$$\iota^*(g) = \arg\max \iota^b \eta^{1-b} \quad \text{s.t. } g = \alpha\iota + \eta. \tag{6}$$

It follows from the first-order (necessary and sufficient) conditions of the problem above that $\iota^*(g) = \frac{bg}{\alpha}$. Hence, productivity is increasing for $\iota \leq \iota^*(g)$ and decreasing for $\iota \geq \iota^*(g)$. The interpretation of this result is straightforward. Given the result of the interview g, firms prefer moderate intelligence levels since a worker whose intelligence is too high must have a low level of perseverance.

But, as a worker must be earning her expected productivity in any separating set, it follows that wages are non-monotone in intelligence (controlling for the interview g). As shown in the previous signaling literature, when the SCP is satisfied, the educational level is increasing in the worker's characteristics. Suppose this is also the case when the SCP is not valid (i.e., suppose that education is increasing in intelligence). Then, firms would offer a higher salary for individuals with intermediate schooling (as those are the most productive workers).[9] But such an allocation cannot be an equilibrium since workers' strategies are not optimal: if they reduce the amount of schooling, their wages rise (and, of course, they obtain a higher utility). Hence, a necessary condition for truth-telling is that education must be increasing in ι until ι^* and decreasing after ι^*.

Notice that this necessary condition for an interior solution follows from the equalization between the marginal benefit from deviating and its marginal cost. Since the marginal benefit consists of the wage differential s_ι and the marginal cost consists of the marginal cost of signaling times the signaling differential $c_y y_\iota$, we get, by computing s_ι and c_y, that

$$y_\iota(\iota, g) = s(\iota)(bg - \alpha\iota), \tag{7}$$

which implies that y must be increasing (decreasing) if $\iota < (>) \iota^*(g)$.

From the local second-order condition, we obtain the usual necessary condition that education must be increasing in the CS_+ region and decreasing in CS_-. Hence, from the first- and second-order conditions of the problem defined in equation (5) we obtain the following lemma, whose proof is presented in the appendix:[10]

[9] More precisely, the wage schedule would be increasing in schooling until $y(\iota^*(g), g)$ and decreasing from that point on.

[10] Lemma 1 can be generalized to C^1 by parts functions. However, we focus on the C^2 by parts case in order to simplify the proof.

Lemma 1 *In any separating set, if a C^2 by parts education and wage profile is truth-telling it must satisfy*

$$y_\iota(\iota, g)(g - 2\alpha\iota) \geq 0. \tag{8}$$

and equation (7).

Corollary 1 *In a separating set, the workers with highest schooling are the most productive ones (not the brightest or the most perseverant ones) and schooling is strictly increasing in productivity.*

Proof. From (7), it follows that

$$y_\iota(\iota, g) > 0 \Longleftrightarrow \iota < \frac{bg}{\alpha} = \iota^*(g). \tag{9}$$

■

Remark 1 *Notice that equation (8) implies that*

$$y_\iota(\iota, g) \geq 0 \Longleftrightarrow \iota \leq \frac{g}{2\alpha}. \tag{10}$$

Generally, equations (9) and (10) cannot hold for all ι except when $b = \frac{1}{2}$. In this case, the firms' technology is identical to the signaling technology. Then, we can treat $\iota\eta$ as a single parameter and we obtain Spence's (1973) model. Moreover, education must be monotone in this (redefined) parameter.

Remark 2 *When $b \neq \frac{1}{2}$, there exists some misalignment between the firm and the worker since the relative intensity of intelligence of the schooling technology is different from that of the firm's technology. Then, if $\min\left\{\frac{bg}{\alpha}; \frac{g}{2\alpha}\right\} \in [\iota_0, \iota_1]$, there must exist some pooling in equilibrium (since the separating set conditions cannot hold for all the interval of parameters).*

1.3.2 Continuous pooling set

Let $p(\iota \mid g)$ denote the density function of ι conditional on the result of the interview g and suppose there exists a non-degenerate closed set I which is a continuous pooling set and such that no closed set $X \supset I$ is a continuous pooling set. Then, $y(\iota, g) = \bar{y}(g)$ for all $\iota \in I$.

The zero-profit condition is

$$w(\bar{y}(g), g) = W(X, g), \tag{11}$$

where $W(X, g) \equiv \int_X f(x, g - \alpha x)\, p(x \mid g) dx$ is the expected productivity of a type-ι worker. Conditions 2 and 3 from Definition 1 are thrivially satisfied in that given set.

1.3.3 Discrete pooling set

A distinct feature of models where the SCP does not hold is the emergence of discrete pooling, where individuals with non-adjacent types receive the same contract [Araujo and Moreira (2000, 2001)]. This feature is a direct consequence of the possibility of non-monotone signals.

As was shown by Araujo and Moreira (2000), a necessary condition for truth-telling in a discrete pooling set is the so-called marginal rate of substitution identity, according to which, if two individuals are (discretely) pooling in a contract, they should have the same marginal rate of substitution. We formally state that result as a lemma:

Lemma 2 *If two regular workers with the same interview result choose the same level of education, then their marginal cost of education must be the same:*

$$\left.\begin{array}{c} y(\iota, g) = y(\tilde{\iota}, g) \\ y_{\iota}(\iota, g) \neq 0 \\ y_{\iota}(\tilde{\iota}, g) \neq 0 \end{array}\right\} \Rightarrow \frac{\partial c(\iota, g, y)}{\partial y} = \frac{\partial c(\tilde{\iota}, g, y)}{\partial y}.$$

Remark 3 *The economic interpretation of Lemma 2 is that if two non-adjacent workers with different marginal costs of education choose the same contract, one of them could benefit from deviating by choosing a different amount of schooling.*

From the equality of the marginal costs of signaling, it follows that if type-(ι, g) worker is in a discrete pooling set, the other worker pooling with her is $(\hat{\iota}, g)$ defined as:

$$\hat{\iota} = \frac{g}{\alpha} - \iota \equiv \gamma(\iota). \tag{12}$$

The following lemma will be important for the extension of the model to the GED exam. It links the productivity of two discretely pooled workers with the relative intensity of cognitive skills in the firms' production function.

Lemma 3 *If two workers are discretely pooled, then the less intelligent one is more productive if the firms' technology is intensive in perseverance ($b < \frac{1}{2}$) and the more intelligent one is more productive if the firms' technology is intensive in intelligence ($b > \frac{1}{2}$).*

Let $P(x)$ denote the density of a type-x individual conditional on x belonging to the pooling-set $\Theta(y(\iota), g)$. Then, if x belongs to a discrete-pooling set, it follows that

$$P(x) \equiv \frac{p(x \mid g)}{p(\iota \mid g) + p(\gamma(\iota) \mid g)}.$$

Furthermore, $P(\iota) + P(\gamma(\iota)) = 1$ for all ι in a discrete-pooling set.

Analogously to Lemma 1, the local first- and second-order conditions from the workers' truth-telling conditions yield the following:

Lemma 4 *If (ι, g) belongs to a discrete pooling set, then if a C^2 by parts education and wage profile is truth-telling, they satisfy:*

$$\begin{aligned} y_\iota(\iota, g) = f(\iota, \eta)\left[P(\iota)(bg - \alpha\iota) + P'(\iota)\right] \\ + \alpha^{1-2b} f(\eta, \iota)\left\{(1 - P(\iota))\left[(1-b)g - \alpha\iota\right] + P'(\iota)\right\}, \end{aligned} \tag{13}$$

$$y_\iota(\iota, g)(g - 2\alpha\iota) \geq 0. \tag{14}$$

Equation (13) displays how discrete pooling distorts an incentive-compatible profile of education. As in the separated case, equation (13) equates the marginal cost with the marginal benefit of education. However, due to the fact that in the discrete pooling case wages are a weighted average of productivities, the marginal benefit of education in a discrete pooling set is a weighted average of marginal productivities.[11]

In the next subsection, we present some comparative statics results as well as the equilibrium selection criterion.

1.3.4 Equilibrium selection and comparative statics

The proposition below presents some comparative statics results. Since education is costly, individuals would only choose to educate if this increases their wages. Thus, incentive-compatibility requires wages to be strictly monotonic.

Proposition 1 *Wages are strictly increasing and concave in the amount of schooling controlling for the interview.*

Notice that productivity is increasing in the result of the interview g. Then, in a separating set, wages must be increasing in g. However, this may not be true in a pooling set: since wages are a weighed average of the productivity of pooled types (where weights are given by the conditional probability of each type), a change in g would also affect the weights attributed to each type. In a discrete pooling set, for example, it follows that[12]

$$\frac{\partial w}{\partial g} = P(\iota) f_\eta(\iota, \eta) + [1 - P(\iota)] f_\eta(\hat{\iota}, \hat{\eta}) + \frac{\partial P(\iota)}{\partial g}[s(\iota) - s(\hat{\iota})].$$

The first and second terms are positive and represent the direct effect: more productive individuals get a higher result in the interview. The last term may be eigher positive or negative and reflects the indirect effect. If the amount of more productive individuals is decreasing in g, then this term is negative.[13] If $\iota \mid g$ is uniformly distributed, for example, then this last term vanishes (since the conditional distribution of ι is constant) implying that wages are increasing in the interview.

[11] Notice that the separating set is a special case of the discrete pooling set where firms are able to infer the workers ability in a pooling set ($P(\iota) = 1$).

[12] The same argument also holds for continuous pooling sets.

[13] Let $s(\iota) > s(\hat{\iota})$. Then, $\frac{\partial w}{\partial g} < 0$ if and only if $\frac{\partial P(\iota)}{\partial g} < -\frac{P(\iota) f_\eta(\iota,\eta) + [1-P(\iota)] f_\eta(\hat{\iota},\hat{\eta})}{s(\iota) - s(\hat{\iota})}$.

The difference between the monotonicity of wages in education (Proposition 1) and the non-monotonicity of wages in the interview stems from the fact that education is an endogenous signal while the interview is an exogenous signal. When a costly signal is endogenous, an agent will not purchase an additional amount of it unless he obtains higher wages by doing so. In contrast, when a signal is exogenous, the agent is unable to distort it. Hence, wages may be non-monotonic in this signal.

As the concept of PBE leads to an indeterminacy of equilibria, it is important to apply a selection criterion in order to pick an equilibrium. Riley (1979) suggested the concept of a reactive equilibrium that chooses only the separating equilibrium in the continuous-type framework. This concept has been widely applied in the signaling literature.

As a fully separating equilibrium does not exist when the single-crossing property does not hold, one must employ a weaker refinement criterion. Araujo and Moreira (2001) proposed the quasi-separability criterion which consists of a slight modification to the concept of reactive equilibrium (both concepts are equivalent when the SCP holds).

Like the reactive equilibrium, the quasi-separable equilibrium seeks a unique equilibrium with the highest degree of separation and which Pareto dominates other signaling equilibria. The following definition introduces the quasi-separability criterion.

Definition 4 *A PBE is quasi-separable if:*

1. *A worker belongs to a pooling set, then there exists a worker with a different type that pools with him such that their marginal cost of schooling must be the same;*

2. *There is no other PBE satisfying condition 1 such that every type obtains less schooling (with strictly less to at least one type).*

The first condition identifies the highest possible degree of separability. The second condition gives the boundary condition which uniquely determines the equilibrium. It consists on a Pareto improvement criterion for selection.

The following proposition can be seen as an evidence that the SCP does not hold. It states that two individuals with different abilities obtaining the same amount of schooling is not consistent with the SCP. Hence, the fact that the empirical evidence documents that workers with different abilities receive the same wages suggests that the SCP is violated.

Proposition 2 *If the pooling set of a quasi-separable equilibrium is non-empty, then the SCP does not hold.*

1.3.5 Characterization of the equilibrium

In this section, we characterize the equilibrium of the model. As the results are more technical than the rest of the paper and are not crucial to any of our results, it can be skipped by readers more interested in the applications of the model.

As in equation 12, we denote by $\gamma(\iota)$ the type with the same marginal cost of signaling as ι. We will focus on the case where $\gamma(\iota_0) \leq \iota_1$ and $b < 1/2$ (the other cases can be studied in a similar fashion).[14] Clearly, as $\gamma(\iota_0) \leq \iota_1$, it follows that $(\gamma(\iota_0), \iota_1]$ must be a separating set in any quasi-separable equilibrium (as no other type can have the same marginal cost of schooling as $\iota \in (\gamma(\iota_0), \iota_1]$). The characterization will be done through a series of lemmata.

Define the indirect utility $U(\hat{\imath}, \iota)$ as the utility received by a type-ι worker who gets the contract designed for type $\hat{\imath}$:

$$U(\hat{\imath}, \iota) \equiv s(\hat{\imath}) - c(\iota, g, y(\hat{\imath}, g)).$$

The first lemma establishes another necessary condition for truth-telling.[15]

Lemma 5 *$U(\hat{\imath}, \iota)$ is continuous at $\hat{\imath} = \iota$ for all $\iota \in [\iota_0, \iota_1]$.*

The basic intuition behind this result is that, as the cost of signaling is continuous, if the indirect utility were discontinuous those individuals in a vicinity of the point of discontinuity could benefit from another type's contract. Hence, it would not be truth-telling.

The continuity of U enables us to determine the boundary condition for the amount of education when changing from discrete pooling sets to separating sets. Notice that when a worker becomes pooled with another type, his expected productivity changes discontinually (as it becomes the average of their productivities). Thus, his wage becomes discontinuous. Hence, the education must be discontinuous in order to preserve the continuity of the indirect utility. This is formally established in the following corollary:

Corollary 2 *(i) Let ι be such that $[\iota, \iota+\varepsilon]$ is a discrete pooling set and $[\iota-\varepsilon, \iota)$ is a separating set, for some $\varepsilon > 0$. The following condition is necessary for truth-telling:*

$$y(\iota) = -\frac{\iota(g-\alpha\iota)[s(\iota) - s(\gamma(\iota))]}{2} + \lim_{x \to \iota_-} y(x). \tag{15}$$

(ii) Let ι be such that $[\iota-\varepsilon, \iota]$ is a discrete pooling set and $(\iota, \iota+\varepsilon]$ is a separating set, for some $\varepsilon > 0$. The following condition is necessary for truth-telling:

$$y(\iota) = -\frac{\iota(g-\alpha\iota)[s(\iota) - s(\gamma(\iota))]}{2} + \lim_{x \to \iota_+} y(x). \tag{16}$$

The second lemma determines the maximal discrete pooling set.

[14] See Araujo, Gottlieb, and Moreira (2004a) for a characterization of the equilibrium in more general models.

[15] Lemma 5 could also be seen as an implication of the Theorem of the Maximum by establishing the continuity of wages in education.

Lemma 6 $[\iota_0, \gamma(\iota_0)]$ *is a discrete pooling set.*

As the set $(\gamma(\iota_0), \iota_1]$ must be separated, it follows from Lemma 6 that the set of types can be partitioned in two intervals: a discrete pooling interval $[\iota_0, \gamma(\iota_0)]$ and a separated interval $(\gamma(\iota_0), \iota_1]$.

The next lemma determines the boundary condition which gives the equilibrium amount of education. It ensures that the individual with the lowest productivity chooses to get no education.

Lemma 7 *In any quasi-separable equilibrium,* $y(\iota_1) = 0$.

The proof basically shows that as ι_1 is separated and is the least productive type, reducing the amount of schooling would never reduce its wages (as no firm would ever expect some type to be less productive than ι_1). But this would also reduce the cost of schooling. Thus, in equilibrium, ι_1 must choose the lowest amount of schooling possible.

The following lemma establishes sufficient conditions for an equilibrium. It was demonstrated by Araujo and Moreira (2000).

Lemma 8 *The differential equations from Lemmas 1, 4 and the boundary conditions from Lemmas 5 and 7 are sufficient for the quasi-separable equilibrium.*

The next proposition is a direct consequence of Lemmas 1, 4, 7, and 8, and Corollary 2.

Proposition 3 *A C^2 by parts education profile is a quasi-separable equilibrium if, and only if it satisfies:*

1. $y_\iota(\iota, g) = s(\iota)(bg - \alpha\iota)$, *for* $\iota > \gamma(\iota_0)$;
2. $y(\iota_1, g) = 0$;
3. $y_\iota(\iota, g) = f(\iota, \eta)[(bg - \alpha\iota) P(\iota) + P'(\iota)] +$
$\alpha^{1-2b} f(\eta, \iota)\{P(\gamma(\iota))[(1-b)g - \alpha\iota] - P'(\gamma(\iota))\}$, *for* $\iota \leq \gamma(\iota_0)$;
4. $y(\gamma(\iota_0), g) = \frac{\iota_0(g - \alpha\iota_0)[s(\iota_0) - s(\gamma(\iota_0))]}{2} + \lim_{x \to \gamma(\iota_0)+} y(x)$.

Proposition 3 is useful as it reduces the problem of obtaining an equilibrium profile of education to that of solving two ordinary differential equations with given boundary conditions. As both differential equations are Lipschitz, it follows that the quasi-separable equilibrium exists and is unique.

The amount of education for separated types is determined from the first equation of Proposition 3 and the boundary condition is given by $y(\iota_1) = 0$. Then, using conditions 3 and 4 from Proposition 3 (a differential equation with a boundary condition), one can calculate the equilibrium amount of education for discrete pooling types.

Notice that item 4 from Proposition 3 implies that education must jump downward at $\gamma(\iota_0)$ since $s(\iota_0) - s(\gamma(\iota_0)) > 0$ (see Lemma 3 and $b < 1/2$).

This follows from the fact that wages are discontinuous: individuals with $\iota \in \left[\frac{g}{2\alpha}, \gamma(\iota_0)\right]$ earn wages higher than their productivity since they are pooled with more productive workers but those with types higher than $\gamma(\iota_0)$ earn their productivity since they are separated. Hence, if education were continuous, indirect utility would be discontinuous. But, as shown in Lemma 5, a discontinuous indirect utility is not incentive-compatible. Thus, the amount of education must jump downward in order to preserve the continuity of the indirect utility function.

The following graphs present the equilibrium amount of education, wages and utility for the case where $b = 0.4$, $g = 10$, $\alpha = 1$, $\iota_0 = 1$, $\iota_1 = 10$, and $\iota \mid g \sim U[\iota_0, \iota_1]$.[16]

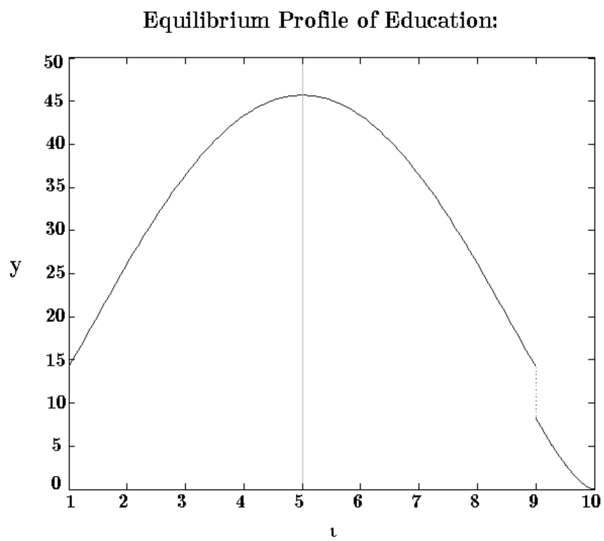


Figure 2

[16] Araujo, Gottlieb, and Moreira (2004b) present the equilibrium profiles of education, wages, and utility for other parameters.

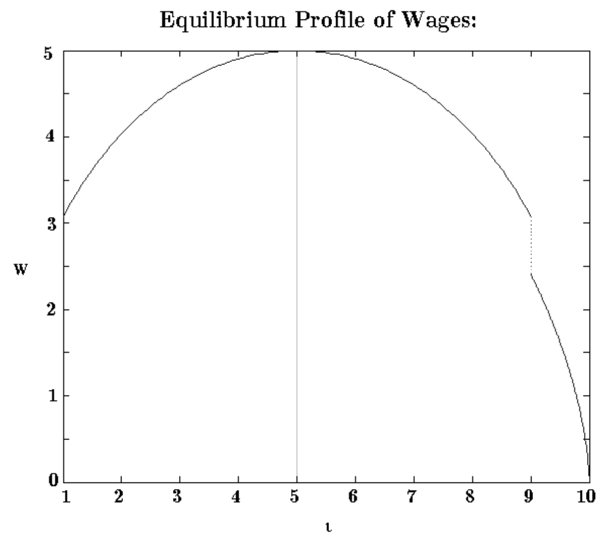


Figure 3

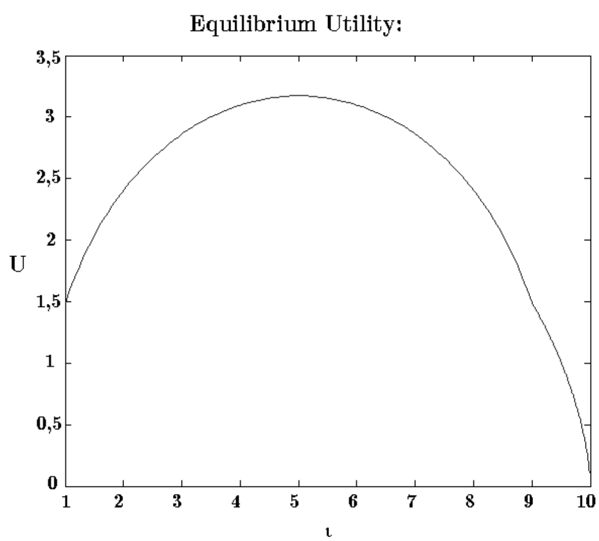


Figure 4

Notice that both education and wages are discontinuous but the utility is countinous in ι. In the graph below, the profile of wages as a function of education is presented. As Proposition 1 shows, wages are strictly increasing and concave in education.

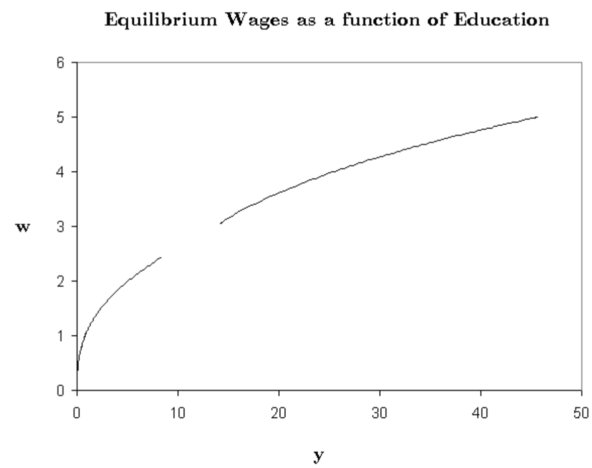


Figure 5

1.4 Countersignaling

In some situations, high-productivity individuals may choose to signal less than those with lower productivity. Clements (2004) argues that many high-quality products are sold in low-quality packages. Moreover, Caves and Greene (1996) found no significant systematic positive correlation between quality and advertising.

O'Neil (2002) argues that the fact that most advanced countries had lower military spendings than those intermediately advanced after World War II occurred due to countersignaling.

Hvide (2003) argues that individuals with intermediate abilities educate more than bright individuals in professions where a license is not required to work.

In their article, Feltovich et. al (2001) present a countersignaling model applied to the labor market. As in our basic framework, firms access some measure of the worker's ability (which is interpreted as the recommendation of a former boss). This signal consists of the sum of the unidimensional ability of the worker and a noise term. Workers may also engage in schooling activity. In equilibrium, low-ability workers and high-ability workers choose not to participate in the signaling game. This occurs since a productive worker has very high probability of receiving a good recommendation and a low-productivity worker has very low probability of receiving good recommendation. Thus, in both cases, the expected benefits from signaling are not sufficiently high. For individuals who choose to participate in the signaling game, the equilibrium is equivalent to the traditional signaling models.

Unlike the model of Feltovich et. al (2001), the uncertainty about productivity comes from the fact that the schooling technology differs from the firms' technology in our model. This misalignment between these two technologies generates an incentive for some higher-productivity workers to educate less. Thus, while in their model the presence of another signal implies that some types may choose not to participate, countersignaling in this model emerges due to incentive reasons.

Orzach et. al (2002) present a model where firms signal product quality through prices and advertising expenditures. Product quality is represented by a parameter that may take two values. Their main conclusion is that modest advertising can be used as a signal of high quality. However, as their model features only two types of firms, they are unable to consider the emergence of non-monotone signals.

In this section, we show how the basic model presented allows us to understand the existence of countersignaling.

First, we present a precise definition of countersignaling.

Definition 5 *A type-(ι, g) worker is countersignaling if*

$$sgn\{y_\iota(\iota, g)\} \neq sgn\{s_\iota(\iota)\}.$$

The definition above states that countersignaling occurs if more productive individuals choose less education than intermediate individuals. With no loss of generality, we can restrict to the case where $b \leq \frac{1}{2}$ (since we can always relabel ι and η).

As shown in Section 1.3.5, education is strictly increasing for $\iota < \frac{g}{2\alpha}$ and strictly decreasing for $\iota > \frac{g}{2\alpha}$. Moreover, as argued in page 8, the productivity of a worker with interview result g is strictly increasing for $\iota < \frac{bg}{\alpha}$ and strictly decreasing for $\iota > \frac{bg}{\alpha}$. Then, the countersignaling interval is $[\frac{bg}{\alpha}, \frac{g}{2\alpha}]$. Hence, countersignaling occurs if, and only if, the schooling technology is not the same as the firms' technology $b \neq \frac{1}{2}$.

Define the distance between the Cobb-Douglas functions $f(\iota, \eta) = \iota^b \eta^{1-b}$ and $\tilde{f}(\iota, \eta) = \iota^{\tilde{b}} \eta^{1-\tilde{b}}$ as $\left| b - \tilde{b} \right|$. Then, the distance from the schooling technology to the firms' technology is given by $\frac{1}{2} - b$. Notice that increasing the distance between the two technologies (i.e., reducing b) strictly increases the countersignaling interval. Thus, we have proved the following:

Proposition 4 *Countersignaling occurs if and only if the schooling and the firms' technologies are not the same (i.e., the SCP does not hold), and the countersignaling interval is strictly increasing in the distance from the schooling technology to the firms' technology.*

This proposition provides an intuitive testable implication. Countersignaling is expected to occur more often in occupations that require a different combination of skills than those required at school. Hence, productive individuals with low educations should be more common among sportsmen or artists than among teachers.

1.5 The GED exam

1.5.1 Empirical evidence

The General Educational Development (GED) is an exam taken by American high school dropouts to certify their equivalence with high school graduates. It started in 1942 as a way to allow veterans without a high school diploma to obtain a secondary school credential. Nowadays, about half of the students who have dropped out of high school and a fifth of those classified as high school graduates by the U.S. Census are GED recipients.

The GED consists of five tests covering mathematics, writing, social studies, science, and literature and arts. Except for the writing part, all other sections are made of multiple choice questions. The costs of acquiring a GED are relatively small. The pecuniary costs range from zero dollars in some states to around fifty in other states and the median study time for the tests is only about twenty hours.

Even though the U.S. Census classifies dropouts who have acquired a GED as ordinary high school graduates, the market does not treat them equally. GED recipients earn lower wages, work less in any year and stay at jobs for shorter periods than high school graduates [Boesel, Alsalam and Smith (1998)].

GED recipients are smarter than other dropouts but exhibit more behavior and self discipline problems and are less able to finish tasks. They turn over jobs at a faster rate and are more likely to fight at school and work, use pot, skip school and participate in robberies. Hence, the GED conveys two pieces of information in one signal. The student who acquires it is bright, but lacks perseverance and self discipline [Cameron and Heckman (1993), Cavallo, Heckman and Hsee (1998), and Heckman and Rubistein (2000)].

Cavallo, Heckman and Hsee (1988) and Heckman and Rubinstein (2001) have shown that when one controls for both cognitive and non-cognitive abilities, there is no difference in earnings between a GED recipient and a dropout who has not taken the exam. As for females, the evidence is the same as that of males, except for those who dropped out because of pregnancy [Carneiro and Heckman (2003)].[17]

As dropouts who have taken the GED are treated in the labor market just like those who have not taken it, any theory that tries to explain this exam must exhibit pooling in equilibrium. Moreover, since GED recipients do not earn higher wages, the signal-earnings relation is not strictly monotone as in the traditional signaling models.

Despite of being the usual assumption in signalling models [e.g. Spence (1973, 1974)] and early human capital models [e.g. Becker (1964)], it is widely accepted that an individual's personal abilities cannot be successfully captured by a scalar of cognitive skills. Cawley et al. (1996), for example, showed that cognitive ability is only a minor predictor of social performance and that many non-cognitive factors are important determinants of blue collar wages.

[17] Tyler, Murnane, and Willett (2000) suggested than the GED does not increase wages except for young white dropouts who are in the margin of passing the exam.

Bowles and Gintis (2001) provided an interesting example of the importance of non-cognitive skills for labor market success. From a survey of 3,000 employers (Bureau of the Census, 1998), they were asked "When you consider hiring a new nonsupervisory or production worker, how important are the following in your decision to hire?". On a scale of 1 to 5, employers ranked "years of schooling" at 2.9, and "scores on tests given by employer" and "academic performance" at 2.5. The non-cognitive skills, "attitude" and "communication skills", were ranked at 4.6 and 4.2, respectively.

Weiss (1988) and Klein, Spady and Weiss (1991) showed that lower quit rates and lower absenteeism account for most of the premium awarded by high school graduates compared to high school dropouts (and *not* higher productivity).

Bowles and Gintis (1976) suggest that employers in low skill markets value docility, dependability and persistence more than cognitive skills. Bowles and Gintis (1998) argue that personality and other affective traits reduce the costs of labor turnover and contract enforcement.

In the Psychology field, the widely accepted five-factor model of personality (referred to as the "Big Five") identifies five dimensions of non-cognitive characteristics: extroversion, conscientiousness, emotional stability, agreeableness, and openness to experience. Personality measures based on this model are good predictors of job performance for a wide range of professions [Barrick and Mount (1991)].

Hogan and Hogan (1989), Barrick and Mount (1991), and Boudreau, Boswell, and Judge (2001) show that personality traits are important predictors of successful employment. Goffin, Rothstein and Johnston (1996) demonstrate that personality characteristics predict job performance better than cognitive skills. Dunafon and Duncan (1998, 1999) document that a series of behavioral characteristics measured when young significantly affect earnings 25 years latter. Edwards (1976) shows that dependability and consistency are more valued by blue collar supervisors than cognitive ability.[18]

1.5.2 The Model

In this subsection, we extend the basic framework to study the effect of the introduction of a pass-or-fail test like the GED. We model the GED as a signal $h(\iota, \eta)$ which only individuals with a sufficiently high combination of characteristics are able to receive. More specifically, denoting by $h(\iota, \eta) = 1$ if an individual passes the exam and $h(\iota, \eta) = 0$ if she fails, we specify the test as

$$h = \begin{cases} 1, \text{ if } \kappa\iota + \eta \geq \overline{g} \\ 0, \text{ if otherwise} \end{cases},$$

where $\overline{g} \in \mathbb{R}_{++}$ is the parameter that represents the minimum combination of skills required to pass the test (passing standards) and κ is the rate of substitu-

[18] There is also significant literature on the importance of non-cognitive skills in business organizations [e.g. Sternberg (1985) and Gardner (1993)], and military organizations [e.g. Laurence (1998)].

tion between intelligence and perseverance.[19] We assume that there is no cost in taking the test.[20]

An employer cannot distinguish a worker who failed the GED exam from a worker who did not take it. Hence, a worker who is able to pass the test will take it as long as her utility is not decreased by holding the certificate.

Controlling for the interview result g, h can be rewritten as

$$h = \begin{cases} 1, \text{ if } (\kappa - \alpha)\iota \geq \overline{g} - g \\ 0, \text{ otherwise.} \end{cases} .$$

According to Heckman, Hsee and Rubinstein (1993), the GED exam is intensive in cognitive skills. Hence, we shall assume that the exam h emphasizes intelligence more than the interview g does:

Assumption 1 $\kappa > \alpha$.

Then, each worker with $\iota \geq \frac{\overline{g}-g}{\kappa-\alpha}$ would be able to pass the test. The graphs below separate the interval $[\iota_0, \iota_1]$ in three regions. The first graph depicts the case where $\frac{\overline{g}-g}{\kappa-\alpha} > \frac{g}{2\alpha}$, while the second represents the case where $\frac{\overline{g}-g}{\kappa-\alpha} < \frac{g}{2\alpha}$.

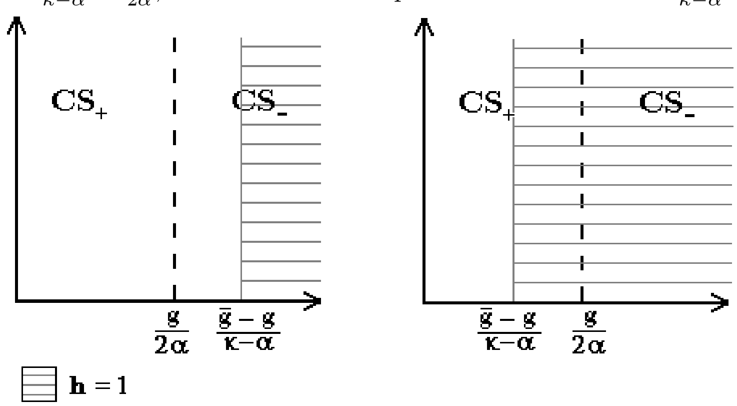


Figure 6

In the left region, workers have low intelligence. Hence, education must be increasing in intelligence (CS_+ region) and the worker is unable to pass the test. In the right side, workers have high intelligence. Thus, education must be decreasing in intelligence (CS_- region) and the worker is able to pass the test.

The region in the middle depends on the sign of $\frac{\overline{g}-g}{\kappa-\alpha} - \frac{g}{2\alpha}$. If $\frac{\overline{g}-g}{\kappa-\alpha} > \frac{g}{2\alpha}$ (first graph), some workers with types in the CS_- region are unable to receive $h = 1$. If $\frac{\overline{g}-g}{\kappa-\alpha} < \frac{g}{2\alpha}$ (second graph), some workers with types in the CS_+ region are able to pass the test.

[19] The assumption that schooling does not affect the possibility of passing the GED is unimportant for our results. As would probably be clear, all results still hold if education entered linearly in the minimum combination of skills.

[20] As the median time studying for the GED exam is 20 hours and the monetary costs range from zero to fifty dollars it seems that the actual costs of taking a GED are very low.

The following proposition is the main result of this section. It states that, as long as the firms' technology is intensive in non-cognitive skills, the introduction of the test does not affect earnings. Thus, we say that in this case the GED is a neutral signal.

Proposition 5 *If the firms' technology is intensive in non-cognitive skills, the introduction of the GED exam does not modify the wage schedule and the profile of education.*

Proof. The result is trivial for a separating set. Assume two workers with $\iota \leq \frac{\overline{g}-g}{\kappa-\alpha} \leq \hat{\imath}$ are pooled in the same contract (otherwise, the signal is not informational). Then type-$\hat{\imath}$ has $h = 1$ (if he chooses to take the exam) and type-ι has $h = 0$. Then, from Lemma 3, the firm would offer a higher salary for the type-ι worker. But this cannot be an equilibrium since the type-$\hat{\imath}$ worker's strategy is not optimal (condition 1).

Thus, any wage schedule such that a type-$\hat{\imath}$ individual earns less than type ι cannot be an equilibrium. Hence, a type-$\hat{\imath}$ individual must earn the same as type ι and is indifferent between taking the GED or not. ■

Remark 4 *Even though the GED does not affect wages, it reveals information about the workers' characteristics. Hence, consistent with Heckman and Rubinstein (2001) and Cavallo, Heckman and Hsee (1998), firms offer the same wages to individuals with low cognitive skills/high non-cognitive skills as to high cognitive skills/low non-cognitive skills individuals.*

Remark 5 *As the result above holds for all $\overline{g} \in \mathbb{R}_{++}$, it follows that, unlike Cavallo, Heckman and Hsee (1998) suggested, an increase in the GED standards $\bar{g}$ would not affect the wages schedule. This implication of the model could be tested as passing standards vary by states and have changed over time. Thus, one could test if the neutrality of the GED is robust to different states and different periods of time.*

Remark 6 *Since the introduction of the GED does not affect the equilibrium amount of education, our model does not support the claim that, when the GED is neutral, it may discourage education [see Cavallo, Heckman and Hsee (1998)].*

Notice that a key assumption for the neutrality of the GED is that the firms' technology is intensive in non-cognitive abilities.[21] The next proposition states that when the firms' technology is intensive in cognitive skills, the GED signal may be non-neutral in equilibrium.

[21] Another assumption which is central to our results is that the GED is not costly. Nevertheless, our results still hold when the GED is costly as long as there exists some external benefits from being a high school graduate. The neutrality of the GED does not depend on the assumption that schooling does not affect the ability to pass on the exam. For example, suppose that an individual would be able to pass on the GED if $\kappa\iota + \eta + \beta y \geq \overline{g}$. Then, the shaded area in Figure 5 would depend on y but if two workers were discretely pooled in a contract, the one who could pass the test would still be the least productive worker.

Proposition 6 *If the firms' technology is intensive in cognitive skills and there are two types pooled in the same contract such that* $\iota_0 \leq \frac{\overline{g}-g}{\kappa-\alpha} \leq \gamma(\iota_0)$*, then the signal is non-neutral: the wage received by a type-*$\hat{\iota}$ *worker will be strictly higher than that of a type-*ι *worker.*

Proof. In this case, the worker with the highest productivity will be $\hat{\iota}$. Hence, signaling $h = 1$ will differentiate him from ι and allows the firm to offer a higher salary. ■

Corollary 3 *A signal h that places more weight to non-cognitive skills* $(\kappa < \alpha)$ *is non-neutral.*

Remark 7 *A way to make the GED exam a non-neutral signal would be to put more emphasis on non-cognitive skills as it would separate two pooled workers with different signs h. Even though it must be significantly harder to design a signal that emphasizes non-cognitive skills, psychologists have developed tests that measure such skills which have been used by companies to screen workers [e.g. Sternberg (1985)].*

Remark 8 *When the GED is non-neutral* $(b > \frac{1}{2})$*, it separates two previously pooled workers. Then, the wage received by the more (less) productive worker increases (decreases). As incentive-compatibility requires that the indirect utility must be continuous, it follows that, in this case, the introduction of the GED increases (decreases) the education obtained by the more (less) productive workers. Hence, another testable implication of the model is that the variance of education should increase when the GED is non-neutral and should remain constant when it is neutral.*

As shown in Propositions 5 and 6, the introduction of an additional signal implements a fully separating equilibrium. It is possible to generalize this result further and show that, in a model where the sign of $c_{y\iota}$ changes n times, it is sufficient to introduce n additional binary signals to implement full separability:

Proposition 7 *Let* n *be the (finite) number of times that* $c_{\theta y}(\theta, y)$ *changes sign.* n *additional binary signals are sufficient to implement a separable equilibrium.*

Proof. See Appendix B. ■

When the SCP holds, Engers and Fernandez (1987) have shown that one signal is sufficient for full separation. Thus, their result is a special case of Proposition 7 when $c_{y\iota}$ does not change signs. This result can be applied to study the optimal design of tests.

1.6 Conclusion

In this paper, we presented a multidimensional signaling model of mixed signals. It was shown that when firms have access to an interview technology, the single-crossing property may not hold. When this is the case, signals are mixed in the sense that they convey two pieces of information.

Two applications of the model were presented. In the first, we analyzed the emergence of countersignaling, where signals are non-monotone in the worker's productivity. It was shown that countersignaling occurs if, and only if, the schooling technology differs from the firm's technology. Moreover, the countersignaling interval is strictly increasing in the distance between the schooling and the firm's technologies. Hence, this phenomenon is expected to be more important in occupations that require more different combination of skills from those required in the schooling process.

In the second application, we introduced the GED exam. It was shown that, consistently with the empirical evidence, a GED recipient has above average cognitive skills and below average non-cognitive skills. When cognitive skills are more valued in the labor market, this new information affects the equilibrium wage. However, when non-cognitive skills are more valued in the labor market than cognitive skills (as suggested by significant empirical evidence), it does not affect the wage schedule.

The main problem with the GED is its focus on cognitive skills. As the firms' main concern is usually on the worker's non-cognitive skills, a non-neutral signal should assign more weight to these kind of skills. Thus, changing its focus to non-cognitive skills would turn it into a non-neutral signal. Moreover, increasing the passing standards with no change of the relative intensity of each skill in the test would not change the equilibrium wages.

Another contribution of this article is to provide an evidence of the importance of the failure of the single-crossing condition in providing intuitive explanations to observed phenomena. As the absence of this property is necessary for the existence of discrete pooling in equilibrium, the fact that an individual with high cognitive ability and low non-cognitive ability receives the same wages as another with low cognitive ability and high non-cognitive ability while an individual with intermediate abilities does not is an evidence of no single-crossing property.

This paper also has a technical interest as it presents a signaling model where the single-crossing condition does not hold. This framework can be employed in a wide variety of multidimensional signaling models and, in particular, other mixed signals. Drazen and Hubrich (2003) presented evidence that interest rates are mixed signals as they show that the government is committed to maintaining the exchange rate, but may also signal weak fundaments. Burtless (1985) provided another example of a mixed signal where a program provided subsidies for hiring severely disadvantaged workers. However, as the program was excessively targeted, the beneficiaries were widely perceived as incapable. Hence, despite of the subsidies, few employers hired the targeted population.

Appendix

A Robustness of the Single-Crossing property

In this section, we characterize the set of functions h and g that satisfy for the single-crossing property (SCP). We shall argue that the results of the model are robust as long as the firms' technology and the schooling technology cannot be ordered according to their technical rates of substitution.

Let the cost of signaling be represented by the twice continuously differentiable function

$$c = \frac{y}{w(\iota, \eta)},$$

which is assumed to be strictly decreasing in ι and η and strictly increasing in y.

The interview technology is represented by the twice continuously differentiable function $g(\iota, \eta)$ which is assumed to be strictly increasing.

From the implicit function theorem, there exists $\varphi(\iota, \bar{g})$ such that

$$\varphi(\iota, \bar{g}) = \eta.$$

Moreover,

$$\varphi_\iota = -\frac{g_\iota}{g_\eta}.$$

Substituting in the cost function, it follows that

$$c = \frac{y}{w(\iota, \varphi(\iota, \bar{g}))}.$$

Hence,

$$c_{y\iota} = -\frac{w_\iota - w_\eta \times \frac{g_\iota}{g_\eta}}{[w(\iota, \varphi(\iota, \bar{g}))]^2}.$$

Thus, the SCP holds if, and only if, $\frac{w_\iota}{w_\eta} - \frac{g_\iota}{g_\eta}$ has a constant sign for all ι, η. Therefore, a necessary and sufficient condition for the SCP to hold is that the technical rates of substitution of the schooling technology and the firms' technology can be ordered.

Suppose, for example, that w and g are both CES functions:[22]

$$\begin{aligned} w &= [\alpha_1 \iota^\rho + \alpha_2 \eta^\rho]^{\frac{1}{\rho}}, \\ g &= [\beta_1 \iota^\gamma + \beta_2 \eta^\gamma]^{\frac{1}{\gamma}}. \end{aligned}$$

Then, the SCP holds if, and only if, $\frac{\eta}{\iota} - \left(\frac{\beta_1}{\alpha_1}\frac{\alpha_2}{\beta_2}\right)^{\frac{1}{\gamma-\rho}}$ has a constant sign for all ι, η.

[22] The functions considered in the model are special cases of the CES when $\gamma = 0$, $\beta_1 = \alpha$, $\beta_2 = 1$, $\rho \to 0$, and $\alpha_1 = \alpha_2 = 1$.

B Number of tests required for full-separability

As shown in Section 1.5, the introduction of the GED implemented full-separability. In this section, we generalize this result for the case where $c_{\theta y}$ changes sign a finite number of times. As special cases, we obtain the result of Section 1.5 as well as Engers and Fernandez's (1987) result that when the single-crossing property holds no additional signal is required.

The following assumption generalize the single-crossing property as well as the double-crossing property of the model presented before.

Assumption A.1 The sign of $c_{\theta y}(\theta, y)$ does not depend on y, and the number times that $c_{\theta y}(\theta, y)$ changes sign is finite.

We denote by n be the number of times that $c_{\theta y}(\theta, y)$ changes sign.

The following assumption is important for the existence of equilibrium.

Assumption A.2 $p, f \in C^1$.

The following proposition states that under Assumptions A.1 and A.2 there always exists a quasi-separable equilibrium.

Proposition 8 *There exists a quasi-separable equilibrium.*

Proof. See Araujo, Gottlieb, and Moreira (2004). ■

We are now able to prove Proposition 7, which states that n additional signals are sufficient to implement a separable equilibrium.

Proof of Proposition 7. From the first condition of Definition 4, for any $y \in \mathbb{R}_+$, there are at most $n+1$ pooled types. Let $k \leq n+1$ be the number of pooled types. With no loss of generality, reorder them as $\theta_1 \leq \theta_2 \leq ... \leq \theta_k$.

Introduce a costless binary signal h_1 such that type-θ_1 is the only worker who is able to obtain $h_1 = 1$ (if more than one type have the same productivity, take any of them). Thus, only the least productive worker is able to pass the h_1 exam. Then, a profile of education and wages such that the utility obtained by type-θ_1 when he takes the test is lower than if he does not take it is not incentive-compatible. Hence, the utility obtained after taking the test must be the same as if the test were not available. Furthermore, if the education obtained by θ_1 changed, the marginal rate of substitution identity would no longer hold.

Thus, it follows that the introduction of the signal h_1 does not change the equilibrium profiles of education and wages but separates type θ_1 from the $\theta_2, ..., \theta_k$ (i.e., the new equilibrium will feature $k-1$ pooled). Repeating the process t times, there will be at most $k-t$ pooled types. Therefore, introducing $k-1$ new signals, it follows that there will be at most 1 type pooling in each contract. ■

C Proofs

Proof of Lemma 1:

Define $U(\hat{\iota}, \iota)$ as the utility received by a type-(ι, g) individual who gets a contract designed for a type $(\hat{\iota}, g)$ individual:

$$U(\hat{\iota}, \iota) \equiv s(\tilde{\iota}) - c(\iota, g, y(\tilde{\iota}, g)).$$

In order to be true-telling, each worker must prefer to announce his own type:

$$U(\iota, \iota) \geq U(\hat{\iota}, \iota), \qquad \forall \hat{\iota}, \iota \in [\iota_0, \iota_1].$$

The following local first- and second-order conditions must be satisfied:

$$\left.\frac{\partial U(\hat{\iota}, \iota)}{\partial \hat{\iota}}\right|_{\hat{\iota}=\iota} = 0, \tag{17}$$

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

The first-order condition yields, for all ι,

$$s_\iota(\iota) - c_y(\iota, g, y(\iota, g))\, y_\iota(\iota, g) = 0 \therefore y_\iota(\iota, g) = s(\iota)(bg - \alpha\iota). \tag{18}$$

Taking the total derivative of the condition above with respect to ι, we get

$$c_{y\iota}(\iota, g, y(\iota, g))\, y_\iota(\iota, g) = s_{\iota\iota}(\iota) - c_{yy}(\iota, g, y(\iota, g))\, y_\iota(\iota, g) \tag{19}$$
$$- c_y(\iota, g, y(\iota, g))\, y_{\iota\iota}(\iota, g).$$

The second-order condition yields

$$s_{\iota\iota}(\iota) - c_{yy}(\iota, g, y(\iota, g))\, y_\iota(\iota, g) - c_y(\iota, g, y(\iota, g))\, y_{\iota\iota}(\iota, g) \leq 0. \tag{20}$$

Substituting (19) in (20), it follows that

$$c_{y\iota}(\iota, g, y(\iota, g))\, y_\iota(\iota, g) \leq 0 \therefore \frac{g - 2\alpha\iota}{(\iota\eta)^2} y_\iota(\iota, g) \geq 0.$$

Thus, $y_\iota(\iota, g)(g - 2\alpha\iota) \geq 0$. ■

Proof of Lemma 2:

Let $\{w(y(\iota, g)), y(\iota, g)\}$ be an incentive-compatible profile of education and wages:

$$\iota \in \arg\max_{\tilde{\iota}} w(y(\tilde{\iota}, g), g) - c(\iota, g, y(\tilde{\iota}, g)),$$

The first-order condition of the problem above yields

$$w_y(y(\iota, g), g) = c_y(\iota, g, y(\iota, g)). \tag{21}$$

Suppose that $y(\iota, g) = y(\tilde{\iota}, g)$ for some regular types $\iota, \tilde{\iota}$. Then, it follows that

$$w_y(y(\iota, g), g) = w_y(y(\tilde{\iota}, g), g).$$

Substituting in equation (21) yields $c_y(\iota, g, y(\iota, g)) = c_y(\tilde{\iota}, g, y(\tilde{\iota}, g))$. ■

Proof of Lemma 3:

Let $\iota > \hat{\iota}$ be two discretely pooled workers and notice that $\alpha\hat{\iota} = \eta$ and $\alpha\iota = \hat{\eta}$. Substituting in the firm's technology yields,

$$f(\iota, g) > f(\hat{\iota}, g) \Longleftrightarrow \iota^b \hat{\iota}^{1-b} > \hat{\iota}^b \iota^{1-b} \Longleftrightarrow 2b > 1.$$

■

Proof of Lemma 4:

From equation (3), the productivity of a type-$\hat{\iota}$ worker can be written as

$$\begin{aligned} s(\hat{\iota}) &= \alpha^{1-2b} (g - \alpha\iota)^b \iota^{1-b} \\ &= \alpha^{1-2b} f(\eta, \iota). \end{aligned}$$

The zero-profit condition is

$$w(y(\iota, g), g) = P(\iota) f(\iota, \eta) + P(\gamma(\iota)) \alpha^{1-2b} f(\eta, \iota),$$

where $P(x) \equiv \frac{p(x|g)}{p(\iota|g)+p(\gamma(\iota)|g)}$.

As in the proof of Lemma 1, define $U(\hat{\iota}, \iota)$ as

$$U(\hat{\iota}, \iota) = P(\hat{\iota}) f(\hat{\iota}, \hat{\eta}) + P(\gamma(\hat{\iota})) \alpha^{1-2b} f(\hat{\eta}, \hat{\iota}) - c(\iota, y(\hat{\iota}, g)),$$

where $\hat{\eta} = g - \alpha\hat{\iota}$.

The truth-telling condition is

$$\iota \in \arg\max_{\hat{\iota}} U(\hat{\iota}, \iota), \qquad \forall \hat{\iota}, \iota \in [\iota_0, \iota_1].$$

The local first-order condition yields, for all ι,

$$U_1(\iota, \iota) = 0. \tag{22}$$

Substituting U in the equation above, we get

$$\begin{aligned} y_\iota(\iota, g) = & f(\iota, \eta) [(bg - \alpha\iota) P(\iota) + P'(\iota)] \\ & + \alpha^{1-2b} f(\eta, \iota) \{P(\gamma(\iota)) [(1-b) g - \alpha\iota] - P'(\gamma(\iota))\}. \end{aligned}$$

Differentiating equation (22) yields

$$U_{11}(\iota, \iota) + U_{12}(\iota, \iota) = 0 \tag{23}$$

The second-order condition is

$$U_{11}(\iota, \iota) \leq 0 \tag{24}$$

Substituting (23) in (24), it follows that

$$U_{12}(\iota,\iota) \geq 0 \therefore (g-2\alpha\iota)\, y_\iota(\iota,g) \geq 0.$$

■

Proof of Proposition 1:

Suppose that wages are not strictly increasing in education.[23] Then, there exist types ι and $\tilde{\iota}$ such that

$$y(\iota,g) > y(\tilde{\iota},g) \text{ and } w(y(\iota,g),g) \leq w(y(\tilde{\iota},g),g).$$

But this is not truth-telling since

$$w(y(\iota,g),g) - \frac{y(\iota,g)}{\iota\eta} < w(y(\tilde{\iota},g),g) - \frac{y(\tilde{\iota},g)}{\iota\eta},$$

concluding the first part of the proof.

In order to establish the concavity of w, consider the indirect mechanism where individuals reveal y instead of θ. Then, the truthfulness condition is

$$y(\theta) \in \arg\max_y w(y) - \frac{y}{\iota(g-\alpha\iota)}.$$

The second-order condition (necessary) is[24]

$$w''(y(\theta)) \leq 0.$$

■

Proof of Proposition 2:

Suppose that type ι belongs to a pooling set. Then, there exists a type $\hat{\iota} = \frac{g}{\alpha} - \iota \neq \iota$ that pools in a contract with ι. Hence, $\iota + \hat{\iota} = \frac{g}{2\alpha}$, implying that ι and $\hat{\iota}$ cannot both belong to CS_+ or CS_-. ■

Proof of Lemma 5:

(i) Suppose that ι is an interior point of either a separating set or a discrete pooling set. Then, as y is continuous (since it solves a differential equation) it follows that:

$$\lim_{x\to\iota_-} U(\iota,x) = \lim_{x\to\iota_+} U(\iota,x) = U(\iota,\iota).$$

[23] This proposition can also be proved using the Chain Rule: since $\frac{\partial w}{\partial \iota} = w_y(y(\iota,g),g)\, y_\iota(\iota,g)$, and

$$sgn\{w_y\} = sgn\{y_\iota\},$$

the result follows.

[24] Another way of demonstrating the monotonicity of w consists of calculating the first-order condition of the indirect mechanism, which yields: $w'(y(\theta)) = \frac{1}{\iota(g-\alpha\iota)} > 0$.

Suppose that $[\iota, \iota+\varepsilon]$ is a discrete pooling set and $[\iota-\varepsilon, \iota)$ is a separating set, for some $\varepsilon > 0$. Clearly, a necessary condition for truth-telling is

$$\lim_{x \to \iota_-} U(x, x) \geq \lim_{x \to \iota_-} U(\iota, x),$$

which means that the last individuals in the separating set would not want to get the contract of the first individual in the discrete pooling set. Then,

$$\begin{aligned} \lim_{x \to \iota_-} U(x, x) &= s(\iota) - \frac{\lim_{x \to \iota_-} y(x)}{\iota(g - \alpha\iota)}, \\ \lim_{x \to \iota_-} U(\iota, x) &= \frac{s(\iota) + s(\gamma_\iota)}{2} - \frac{y(\iota)}{\iota(g - \alpha\iota)}. \end{aligned}$$

Thus, the inequality can be written as

$$y(\iota) \geq \lim_{x \to \iota_-} y(x) - \frac{\iota(g - \alpha\iota)[s(\iota) - s(\gamma_\iota)]}{2}.$$

Another necessary condition for truth-telling is

$$U(\iota, \iota) \geq \lim_{x \to \iota_-} U(x, \iota),$$

which states that the first individual in the discrete pooling set would not want to get the contract of the last individuals in the separating set.

Expanding the indirect utility functions, it follows that

$$\begin{aligned} U(\iota, \iota) &= \frac{s(\iota) + s(\gamma_\iota)}{2} - \frac{y(\iota)}{\iota(g - \alpha\iota)}, \\ \lim_{x \to \iota_-} U(x, \iota) &= s(\iota) - \frac{\lim_{x \to \iota_-} y(x)}{\iota(g - \alpha\iota)}, \end{aligned}$$

implying in

$$\lim_{x \to \iota_-} y(x) - \frac{\iota(g - \alpha\iota)[s(\iota) - s(\gamma_\iota)]}{2} \geq y(\iota).$$

Thus, from these two necessary conditions, we obtain:

$$y(\iota) = -\frac{\iota(g - \alpha\iota)[s(\iota) - s(\gamma_\iota)]}{2} + \lim_{x \to \iota_-} y(x). \tag{25}$$

Substituting in the indirect utility function, it follows that $U(\iota, \iota) = \lim_{x \to \iota} U(x, \iota)$. Analogously, if $[\iota-\varepsilon, \iota]$ is a discrete pooling set and $(\iota, \iota+\varepsilon]$ is a separating set for some $\varepsilon > 0$, then

$$\begin{aligned} y(\iota) &= -\frac{\iota(g - \alpha\iota)[s(\iota) - s(\gamma_\iota)]}{2} + \lim_{x \to \iota_+} y(x), \\ U(\iota, \iota) &= \lim_{x \to \iota} U(x, \iota). \end{aligned}$$

■

Proof of Lemma 6:

From Remark 2, it follows that some types between $\frac{bg}{\alpha}$ and $\frac{g}{2\alpha}$ must be discretely pooled (since there is no continuous pooling in a quasi-separable equilibrium). Assume that some type in $[\iota_0, \gamma(\iota_0)]$ is separated. Then, there must be a $\iota \in \left[\iota_0, \frac{g}{2\alpha}\right]$ such that $\left[\iota, \frac{g}{2\alpha}\right]$ is a discrete pooling set and $[\iota - \varepsilon, \iota)$ is a separated set for $\varepsilon > 0$. From equation 15, it follows that $y(\iota) < \lim_{x \to \iota_0 -} y(x)$ (i.e., y jumps upward when the types become separated). But this is not truth-telling because the marginal cost of education is lower for $\iota + \varepsilon$ than for $\iota - \varepsilon$ for ε sufficiently small (thus, a type-$(\iota + \varepsilon)$ individual would always prefer to get the type-$(\iota - \varepsilon)$ individual's contract). ■

Proof of Lemma 7:

As $\gamma(\iota_1) < \iota_0$, ι_1 is separated. Suppose a type ι_1 worker chooses some strictly positive education $\tilde{y} > 0$. Then, according to equation (4), this worker's wages must be $s(\iota_1)$ in any separating equilibrium (which is the lowest wage since ι_1 is the least productive type). However, she would receive a wage of at least $s(\iota_1)$ if she chose $y = 0$. As $y = 0$ implies in a lower signaling cost and does not reduce her utility, she would be strictly better off by doing so. ■

Chapter 2

Should Educational Policies be Regressive?

Abstract

In this paper, we study the optimal educational policies in an asymmetric information framework. It is shown that when the government is able to transfer wealth between generations, regressive policies, as proposed by De Fraja (2002), are no longer optimal. The optimal educational policy can be decentralized through 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). Furthermore, when the utility function is not quasi-linear, education may not be monotonic in ability.

2.1 Introduction

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

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

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

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

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

[2] This kind of intergenerational transfers is also available in pay-as-you-go social-security

assume that governments are able to transfer between generations.

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

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

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

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

Nevertheless, since the utility function is assumed to be linear in bequests, the optimal policy implies in a large inequality of bequests. We show that when the utility function is concave or when individuals face a choice between working or educating when young, input-regressive policies may no longer be incentive-compatible. Then, the first-best welfare will not be achievable through pigouvian taxes and credit provision.

The non-implementability of input regressive policies result is fairly general as is does not depend on the specific welfare function or in the instruments available to the government. Hence, it also applies to the Rawlsian welfare function case [Fleurbaey et. at. (2002)] and to political economy models. Thus, unlike De Fraja (2002) and Fleurbaey et. al. (2002), it may be the case that education should not be monotonic in ability and monotonic schemes may not even be implementable.

The remainder of the paper is organized as follows. Section 2.2 lays out the framework of the model. The basic structure is the same as De Fraja (2002). In section 2.3 we present the laissez-faire equilibrium. In section 2.4, the government-intervention solution is presented: the first-best equilibrium is characterized (2.4.1), the second-best equilibrium is presented (2.4.2), and the case where there is possibility of default is analyzed (2.4.3). Then, subsection 2.4.4 presents the decentralized equilibrium. Section 2.5 studies the cases where parents are risk-adverse in the wealth left to their children (2.5.1) and where young individuals may work (2.5.2). Section 2.6 summarizes the main results of the paper.

---

systems where young generations contribute for the benefits of the older generations.

2.2 The Basic Framework

We consider an overlapping generations model where individuals live for 2 periods. In the first one (childhood), the individual receives an education and a bequest. In the second period (adulthood), she works, has a child, consumes, and provides an education and a bequest for her daughter. Each household consists of a parent and a child. There is a continuum of households with measure normalized to 1 in every period. As we will focus our attention on steady-state equilibria, time subscripts will be omitted.

Each individual's utility function is

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

where c is her consumption and x is the amount of monetary resources available to the child.[3]

Assumption. $u \in C^2$ satisfies

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

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

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

There are two ways of transferring wealth to the child: bequest t and higher future wages (through education e). We normalize the interest rate paid on bequests to 1. Education is transformed in future wages through the household-production technology $y(\theta, e; E)$, where $\theta \in [\theta_0, \theta_1]$ is each child's productivity parameter, e is the amount of education and E is the general level of education.[4]

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

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

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

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

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

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

Then, the household's budget constraint is

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

Assumption. $y \in C^2$ satisfies

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

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

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

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

The total level of education is defined as the sum of every individual's education:[6]

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

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

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

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

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

2.3 The Laissez-Faire Equilibrium

The imperfection of educational credit markets was studied, among others, by Becker (1964) and Schultz (1963). It is usually argued that investment in human capital is risky, nondiversifiable, and hard to collateralize, implying that private 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 parent's problem consists of determining the optimal amount of consumption, and education and bequest left to her daughter. Hence, the parent's problem is

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

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

1. $\{c(\theta,Y), e(\theta,Y), t(\theta,Y)\}$ *solve the parent's problem for each* $(\theta,Y)\in[\theta_0,\theta_1]\times\Gamma$ *given the educational level* E*, and*

2. *the educational level is the sum of every individual's education:* $E=\int_{\theta_0}^{\theta_1}\int_{Y\in\Gamma} e(\theta,Y)\,\phi(\theta)\,h(Y)\,dY d\theta$.

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

Define $e^u(\theta,Y)$ and $e^c(\theta,Y)$ 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 2 *The laissez-faire competitive equilibrium is* $\{c(\theta,Y), e(\theta,Y), t(\theta,Y);\ \theta\in[\theta_0,\theta_1],\ Y\in\Gamma\}$*, such that*

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

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

[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 and uniqueness of e^u and e^c is demonstrated in the Appendix.

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

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

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

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

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

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

■

Remark 3 *As can be seen in the proof above, if* $Y < c^* + ke^u(\theta,Y)$*, the parent's decisions are constrained since she would prefer to leave negative bequests but is not allowed to. So, she partially reduces her consumption and partially reduces her daughter's education, and leaves no bequests. Since education is increasing in* θ*, households with sufficiently bright children (high* θ*) or low wealth* Y *are constrained.*[9] *Thus the laissez-faire equilibrium is characterized by inequality of opportunities in the sense that individuals with the same ability receive different amounts of education. Moreover, the marginal productivity of children from constrained parents is higher than those of children from unconstrained parents.*

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

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

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

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

2.4 The Government-Intervention Solution

2.4.1 The First-Best Equilibrium

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

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

Definition 4 *A first-best equilibrium is a profile of education and bequests* $\{e(\theta, Y), t(\theta, Y);\ \theta \in [\theta_0, \theta_1],\ Y \in \Gamma\}$ *solving*

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

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

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

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

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

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

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

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

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

The assumption above guarantees that there are enough resources so that e^* and c^* are feasible in a perfect-information economy.[10]

Proposition 5 *The first-best allocations are*

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

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

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

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

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

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

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

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

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

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

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

Remark 6 *As the first-best amount of education $e^*\left(\theta,Y\right)$ is independent of the household's wealth Y, it follows that the optimal educational policy is characterized by equality of opportunities. Since efficiency requires that marginal productivity of education must be equalized for all individuals, the amount of education received by an individual should depend only on her ability. Therefore, we shall denote $e^*\left(\theta,Y\right)$ as $e^*\left(\theta\right)$ in order to emphasize that it does not depend on Y. Notice that as c^* is independent of Y, the optimal consumption level is also independent of wealth.*

Remark 7 *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).*[11] *Moreover, equality of opportunities in education implies that the first-best equilibrium is output-regressive (i.e., individuals with higher ability obtain more utility than lower-ability individuals). Hence, the first-best equilibrium is characterized by an inequality of outcomes.*

Remark 8 *Notice that the presence of positive externalities implies an inefficiently low amount of education provided in the laissez-faire equilibrium even for unconstrained households (since y is strictly concave in e).*

2.4.2 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 $\tau\left(Y\right)$. An education schedule consists of an offer of education $e\left(\theta,Y\right)$, an up-front fee $f\left(\theta,Y\right)$ and a deferred payment

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

$m(\theta, Y)$. Since $f(\theta, Y)$ and $m(\theta, Y)$ may be positive or negative, the government is able to offer loans to students.[12] We assume that the parent's wealth Y is observable but the daughter's ability θ is private information.

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

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

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

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

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

Lemma 9 *A C^2 by parts policy $\{\tau(Y), f(\theta, Y), m(\theta, Y), e(\theta, Y);\ \theta \in [\theta_0, \theta_1],$ $Y \in \Gamma\}$ is incentive-compatible if, and only if, it satisfies*

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

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

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

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

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

2.4.2.1 The Government Budget Constraint

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

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

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

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

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

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

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

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

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

Remark 10 *As deferred payments and education are the only channels of transferring wealth between generations, it follows that imposing that the aggregate amount of deferred payments be non-negative is equivalent to imposing that education is the only means of transferring wealth between generations. However, as public educational systems are usually financed through taxes paid by other generations, it may seem reasonable to allow taxes to be used in order to transfer resources between different generations.*

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

Definition 11 *A second-best equilibrium is a policy* $\{\tau(Y), f(\theta, Y), m(\theta, Y), e(\theta, Y);\ \theta \in [\theta_0, \theta_1],\ Y \in \Gamma\}$ *solving*

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

s.t.(3),(6),(7),(8),(9),(10).

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

Proposition 12 *The optimal educational policy implements the first-best amount of education and consumption* $\{e^*(\theta), c^*;\ \theta \in [\theta_0, \theta_1],\ Y \in \Gamma\}$ *and achieves first-best welfare.*

The basic intuition behind this result is that when the government raises the 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 constraint can be implemented at no cost. Moreover, since the individuals are risk-neutral, any redistribution of wealth does not change the total welfare.

Remark 13 *As shown in Proposition 12, when transfers between generations are allowed, the optimal educational policy provides equality of opportunities in education (since $e^*(\theta)$ does not depend on Y). Furthermore, as was shown in the previous section, the efficient amount of education is higher than the amount provided in the laissez-faire equilibrium. Hence, contrary to the results obtained by De Fraja (2002), the amount of education and consumption does not depend on each parent's wealth.*

2.4.3 The economy with default

In the model presented, the returns to education are deterministic. In reality, however, individuals can neither be sure about finishing their education successfully nor about their future returns after a successful conclusion. Hence, educational returns display a very high variation since students may not graduate or not find a job.[13]

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

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

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

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

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

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

$$U(\theta, Y) = u(Y - \tau(Y) - f(\theta, Y)) +$$
$$y(\theta, e(\theta, Y); E) - \psi(\theta, Y) m(\theta, Y) + t(\theta, Y), \tag{11}$$

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

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

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

Then, the incentive-compatibility constraint is

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

As in Lemma 9, the incentive-compatibility constraint can be written as the local first- and second-order conditions.

Lemma 14 *A C^2 by parts policy $\{\tau(Y), f(\theta, Y), m(\theta, Y), e(\theta, Y); \ \theta \in [\theta_0, \theta_1], Y \in \Gamma\}$ is incentive-compatible if and only if it satisfies*

$$U_\theta(\theta, Y) = y_\theta(\theta, e(\theta, Y); E) - \psi_\theta(\theta, Y) m(\theta, Y), \tag{12}$$

$$y_{e\theta}(\theta, e(\theta, Y); E) e_\theta(\theta, Y) - \psi_\theta(\theta, Y) m_\theta(\theta, Y) \geq 0. \tag{13}$$

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

$$\bar{E}[\tau(Y) + f(\theta, Y) + \psi(\theta, Y) m(\theta, Y) - t(\theta, Y)] \geq \bar{E}[ke(\theta, Y)]. \tag{14}$$

Then, the government faces the following problem:

$$\begin{gathered} \max_{\{e,\tau,f,m,E\}} \int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) U(\theta, Y) \phi(\theta) \, dY d\theta \\ s.t. \text{ (3), (9), (11), (12), (13), (14),} \\ m(\theta, Y) \geq 0, \\ t(\theta, Y) \geq 0. \end{gathered}$$

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

Proposition 15 *The optimal educational policy in the economy with default implements the first-best amount of education and consumption $\{e^*(\theta, Y), c^*; \ \theta \in [\theta_0, \theta_1], \ Y \in \Gamma\}$ and achieves first-best welfare.*

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

Remark 16 *If ψ is interpreted as the probability that the labor income of a type-θ individual is higher than a threshold Y^*, then the optimal educational policy can also be implemented through an income-contingent type of payment since the deferred payment is charged only if the realization of income exceeds this threshold.*[14]

Notice that as preferences are quasi-linear, the government does not wish to redistribute resources. Hence, it is similar to the regulation model when the shadow-cost of public funds is zero. In the next section, we show that the first-best welfare can also be implemented when wealth is unobservable through Pigouvian taxes and public provision of credit. This implementation is desirable due to its simplicity and informational advantage.

2.4.4 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.[15] As Carlton and Loury (1980) show, efficiency may require an additional lump-sum tax-subsidy scheme. In this section, we show that the optimal mechanism can be decentralized through appropriate Pigouvian taxes and the provision of credit at the market interest rate. Moreover, the decentralized scheme does not require knowledge of household's wealth.

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

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

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

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

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

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

$$\int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) \left[\left(k - \hat{k} \right) e(\theta, Y) - \tau \right] \phi(\theta)\, dY d\theta \leq 0. \tag{15}$$

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

Substituting the definitions of $\hat{k}$ and t in the parent's budget constraint (10), it follows that the total amount of consumption, loans repaid, and taxes must be equal to the household's wealth:

$$Y = c + t(\theta, Y) + \tau + \hat{k} e(\theta, Y). \tag{16}$$

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

$$\max_{\{e(\theta,Y), t(\theta,Y)\}} u\left(Y - t(\theta, Y) - \tau - \hat{k} e(\theta, Y) \right) + y(\theta, e(\theta, Y); E) + t(\theta, Y) \tag{17}$$
$$s.t.\ E = \bar{E}[e]$$

As there are no restrictions on t, the solution must be such that the marginal utility of consumption is equal to the marginal utility of wealth left to the daughter. Hence, each parent must be consuming c^*. Moreover, the marginal benefit of education y_e must be equal to its marginal cost $\hat{k}$. This result is stated formally in the following lemma:[16]

Lemma 17 *The solution to the mother's problem (17) is* $\{c^P(\theta, Y), e^P(\theta, Y), t^P(\theta, Y);\ \theta \in [\theta_0, \theta_1],\ Y \in \Gamma\}$, *such that:*

$$\begin{aligned} c^P(\theta, Y) &= c^*, \\ \hat{k} &= y_e\left(\theta, e^P(\theta, Y); E \right), \\ t^P(\theta, Y) &= Y - c^* - \tau - \hat{k} e^P(\theta, Y). \end{aligned} \tag{18}$$

Proof. The result follows from the first-order conditions of the household's problem (17). ■

Now, we are ready to show that the first-best solution can be reached through a suitable choice of τ and $\hat{k}$. As in Carlton and Loury (1980), the efficient allocation can be reached in this context through Pigouvian subsidies and a lump-sum

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

tax. This result can be seen as an application of the so-called 'Principle of Targeting' according to which externalities should be corrected by targeting its source directly.

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

Proposition 18 *There exists a second-best equilibrium where the price of education $\hat{k}$ and the tax τ are both constant in (θ, Y). Moreover, this equilibrium implements first-best welfare.*

Proof. Set $\hat{k}$ as

$$\hat{k} = k - \bar{E}\left[y_E(\theta, e^*, \bar{E}[e^*])\right]. \tag{19}$$

Substituting in the first-order conditions of the household's problem (18), we get $e^P(\theta, Y) = e^*(\theta, Y)$.

Set τ as

$$\tau = \bar{E}[e^*]\ \bar{E}\left[y_E(\theta, e^*, \bar{E}[e^*])\right]. \tag{20}$$

Then, it follows that

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

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

From equation (19), it follows that

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

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

$$\bar{E}\left[\left(k - \hat{k}\right) e^*\right] = \bar{E}\left[y_E(\theta, e^*, \bar{E}[e^*])\right]\ \bar{E}[e^*]. \tag{21}$$

Hence, equations (20) and (21) imply that $\bar{E}\left[\left(k - \hat{k}\right) e^* - \tau\right] = 0$. It follows that the government's budget constraint (15) is satisfied. ■

Remark 19 *As $\bar{E}\left[t^P\right] = \bar{E}[t^*] > 0$, the government transfers resources from older individuals to younger individuals (who repay when older). Thus, this mechanism does not satisfy the government budget constraint imposed in De Fraja (2002) which states that deferred payments can not be used to finance other expenses (i.e., $\bar{E}[t] \leq 0$).*

Notice that since τ and $\hat{k}$ do not depend on Y, the first-best welfare can also be implemented in an environment where income is not observable.

Define the household's financial contribution as

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

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

Let x^P denote the wealth left to the daughter:

$$x^P(\theta, Y) \equiv y\left(\theta, e^P(\theta); E\right) + t^P(Y, \theta).$$

Then, it follows that the marginal propensity to bequeath under the Pigouvian scheme is equal to one (i.e., $x_Y^P = 1$). In other words, every additional amount of wealth is left to the future generation. Furthermore, more able individuals receive more wealth through education in the Pigouvian scheme (since $x_\theta^P = y_\theta > 0$). Therefore, the optimal policy generates large inequalities of wealth left to the future generation. This follows from the quasi-linearity of the utility function. In the next section, we study how the results change when the utility function is not quasi-linear.

2.5 Extensions

2.5.1 Preference for Redistribution

Thorough the previous sections, it has been assumed 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 preference for redistribution of bequests. Then, as was shown in subsection.2.4.4, the optimal policy implements first-best welfare but generates large inequalities of wealth left to the future generation.

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

$$U = u(c) + v(x).$$

This assumption implies that parents are risk-averse in the wealth of their children and a utilitarian government has preference for redistribution.

Substituting the parent's budget constraint in her utility function, we get

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

Then, it follows that

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

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

Notice that the sign of $U_{e\theta}$ 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 ($-r_A y_e y_\theta < 0$) concerns equity: an increasing profile of education generates more wealth to those with lower marginal utility. Then, the sign of $U_{e\theta}$ will depend on the preference for redistribution (captured by the risk-aversion coefficient r_A).

The following lemma presents a necessary condition for incentive-compatibility. When preference for redistribution is sufficiently small, an incentive-compatible policy is input-regressive ($e_\theta > 0$) as in the quasi-linear environment. However, incentive-compatible policies are generally not input-regressive (in fact, they may even be input-progressive).

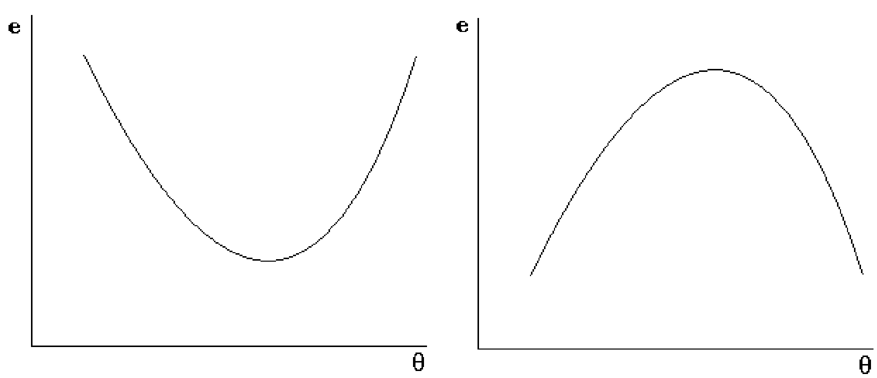
Lemma 20 *An incentive-compatible C^2 by parts policy $\{\tau(Y), f(\theta,Y), m(\theta,Y), e(\theta,Y); \theta \in [\theta_0,\theta_1], Y \in \Gamma\}$ satisfies*

$$U_{e\theta}(\theta,Y)\, e_\theta(\theta,Y) \geq 0,$$

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

Thus, if the concern for redistribution is sufficiently high, input-regressive policies are no longer incentive-compatible. Indeed, incentive-compatible educational policies may be even non-monotonic. This would be the case when the sign of $U_{e\theta}$ changes (i.e., the single-crossing property does not hold).[17]

Suppose, for example, that parents have constant absolute degree of risk-aversion $r_A > 0$. Then, if the sign of $U_{e\theta}$ changes and $\partial\left(\frac{y_{e\theta}}{y_e y_\theta}\right)/\partial\theta > (<)0$, an incentive-compatible educational policy must be U-shaped (bell-shaped) as shown in the following figure:



[17] Araujo and Moreira (2000) present a method for solving screening models where the single-crossing property does not hold.

2.5.1.1 The First-Best Equilibrium

The problem faced by a utilitarian government when ability is observable is

$$\max_{\{e,t\}} u\left(Y - t - ke\right) + v\left(y\left(\theta, e; E\right) + t\right)$$
$$s.a. \quad E = E\left[e\right]$$

The solution to this problem must be such that the marginal cost of bequest must be equal to its marginal benefits for each individual:

$$u'\left(Y - t^{1b} - ke^{1b}\right) = v'\left(y\left(\theta, e^{1b}; E\right) + t^{1b}\right). \tag{22}$$

Moreover, as in the quasi-linear case, the marginal cost of education must be equal to its private marginal benefit plus its social marginal benefit:

$$k = y_e\left(\theta, e^{1b}; E\right) + \frac{E\left[y_E\right]}{u'\left(Y - t^{1b} - ke^{1b}\right)}. \tag{23}$$

Denote c^{1b} as the consumption associated with t^{1b} and e^{1b}:

$$c^{1b} \equiv Y - t^{1b} - ke^{1b}. \tag{24}$$

Then, the results above may be stated as in the following proposition:

Proposition 21 *The first-best allocations are*

$$\left\{c^{1b}\left(\theta, Y\right), e^{1b}\left(\theta, Y\right), t^{1b}\left(\theta, Y\right); \theta \in \left[\theta_0, \theta_1\right], Y \in \Gamma\right\}.$$

From equations (22) and (23), it follows that

$$\frac{de}{dt} = -\frac{k - y_e}{k^2 - ky_e - \frac{y_{ee}}{r_A(u)}} \tag{25}$$

$$= \frac{E\left[y_E\right]}{\frac{y_{ee}}{r_A(u)}u'\left(Y - t^{1b} - ke^{1b}\right) - kE\left[y_E\right]} < 0. \tag{26}$$

Hence, education and bequests are substitutes. From equation (24), it follows that a necessary condition for consumption to be independent of ability is $\frac{de}{dt} = -\frac{1}{k}$ for all Y. Then, equation (25) would imply in

$$\frac{y_{ee}}{r_A(u)} = 0,$$

which contradicts our assumptions about the second derivatives of u and y. Therefore, we have proved that, unlike in the quasi-linear case, the amount of first-best consumption is not constant in θ. The following lemma summarizes this result:

Lemma 22 $c^{1b}\left(\theta, Y\right)$ *is not constant in* θ.

2.5.1.2 Impossibility of a Decentralized Equilibrium

In this subsection, we show that when parents are risk-averse in the wealth left to their children, the first best cannot be achieved in a decentralized equilibrium. The parents' problem is:

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

The solution to this problem will be such that the marginal cost of bequests is equal to its marginal benefits and the marginal cost of education is be equal to its (private) marginal benefits:

$$\begin{aligned} u'\left(Y - t - \hat{k}e\right) &= v'\left(y\left(\theta, e; E\right) + t\right), \\ u'\left(Y - t - \hat{k}e\right)\hat{k} &= v'\left(y\left(\theta, e; E\right) + t\right) y_e\left(\theta, e; E\right). \end{aligned}$$

Therefore, the equilibrium is characterized by the following equations

$$y_e\left(\theta, e^P; E\right) = \hat{k}, \tag{27}$$

$$u'\left(Y - t^P - \hat{k}e^P\right) = v'\left(y\left(\theta, e^P; E\right) + t^P\right). \tag{28}$$

and the overall level of education $E = E\left[e^P\right]$.

In order to implement the first-best amount of education, it is necessary to set $\hat{k}$ (independent of θ) as

$$k - \frac{E\left[y_E\right]}{u'\left(Y - t^{1b} - ke^{1b}\right)}.$$

But this would only be possible if the amount of first-best consumption does not depend on θ. However, as shown in Lemma 22, this is not the case. Hence, we obtain the following proposition:

Proposition 23 *When parents are risk-averse in the wealth left to the daughters, the first-best amount of welfare cannot be reached through credit loans and Pigouvian taxes.*

The intuition for this result is that when the inequality of wealth matters, a policy which does not depend on ability is unable to redistribute appropriately. Then, the Pigouvian mechanism generates an inefficient equilibrium since it does not allow for redistribution of wealth left to the children.

2.5.2 Type-dependent cost of education

Through this section, we have shown two main result for the case where parents are risk-averse: (i) input-regressive policies may no longer be incentive-compatible and (ii) optimal policies cannot be implemented through Pigouvian taxes and loans.

Implicit in the specification of the model presented in this paper is the assumption that the unit cost of education k does not depend on the ability of the individual. However, if individuals are allowed to work when they are not studying, this assumption may not be compelling (since more able individuals would receive higher wages and, thus, would have higher outside options). In this subsection, we show that the input-regressive policies may no longer be incentive-compatible in the quasi-linear environment when individuals are able to work when they are not studying.

Suppose that a type-θ individual receives an hourly wage of $\alpha + \beta\theta$ and let $w \equiv 1 - e$ be the number of hours worked. We interpret α as the part of the job that does not depend on ability and β as the intensity of ability. Then, her parent's budget constraint becomes

$$Y + (1 - e)(\alpha + \beta\theta) = c + ke + t.$$

$$c = Y + \alpha + \beta\theta - t - e(\alpha + \beta\theta + k)$$

Substituting in the quasi-linear utility function yields

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

Then, it follows that

$$U_{e\theta} = y_{e\theta} - \beta\left[u'(c) - (1 - e)(\alpha + \beta\theta + k)u''(c)\right] \tag{29}$$

Once again we obtain an ambiguous sign for $U_{e\theta}$. However, in this case, the indeterminacy arises only for efficiency reasons. An increasing profile of education has the positive effect of benefiting more those with higher marginal productivity of education but has the negative effect of taking more productive individuals more time away from work. Then, as shown in Lemma 20, an incentive-compatible policy may be either increasing, decreasing, or non-monotone, depending on the absolute degree of risk aversion and the magnitude of $y_{e\theta}$.

In this case, the determination of whether an incentive-compatible policy should be input-regressive depends on the level of education studied. It is usually argued that the type of jobs available to very young individuals does not depend much on ability (low β). Then, equation (29) would imply that education for these individuals should be input-regressive since the positive effect exceeds the negative effect. However, ability is clearly an important component in wages of older individuals. Therefore, input-regressive policies may not be incentive-compatible in the case of higher levels of education.

2.6 Conclusion

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

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

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

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

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

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

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

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

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

When parents are risk adverse in the bequests left to their children or when young individuals may work, input-regressive policies may not be incentive-compatible. Moreover, in the former case, the optimal policy cannot be decentralized through Pigouvian taxes and student loans.

A Appendix

2.1.1 Proof of Lemma 9:

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

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

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

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

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

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

Differentiating the first-order condition, we get

$$\begin{aligned}-u'\left(Y-\tau\left(Y\right)-f\left(\theta,Y\right)\right)f_{\theta\theta}\left(\theta,Y\right)+u''\left(Y-\tau\left(Y\right)-f\left(\theta,Y\right)\right)\left[f_\theta\left(\theta,Y\right)\right]^2\\+y_{e\theta}\left(\theta,e\left(\theta,Y\right);E\right)e_\theta\left(\theta,Y\right)+y_{ee}\left(\theta,e\left(\theta,Y\right);E\right)\left[e_\theta\left(\theta,Y\right)\right]^2\\+y_e\left(\theta,e\left(\theta,Y\right);E\right)e_{\theta\theta}\left(\theta,Y\right)-m_{\theta\theta}\left(\theta,Y\right)=0.\end{aligned}\tag{30}$$

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

$$\begin{aligned}-u'\left(Y-\tau\left(Y\right)-f\left(\theta,Y\right)\right)f_{\theta\theta}\left(\theta,Y\right)+u''\left(Y-\tau\left(Y\right)-f\left(\theta,Y\right)\right)\left[f_\theta\left(\theta,Y\right)\right]^2\\+y_{ee}\left(\theta,e\left(\theta,Y\right);E\right)\left[e_\theta\left(\theta,Y\right)\right]^2+y_e\left(\theta,e\left(\theta,Y\right);E\right)e_{\theta\theta}\left(\theta,Y\right)-m_{\theta\theta}\left(\theta,Y\right)\leq 0.\end{aligned}\tag{31}$$

Substituting (30) in equation (31), we obtain

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

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

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 (6), yields

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

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

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

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

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

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

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

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

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

Lemma 20 can be demonstrated in a similar way.

2.1.2 Proof of Proposition 15:

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

$$\begin{aligned} \dot{S}(\theta) &= \int_{Y \in \Gamma} h(Y) e(\theta, Y) \phi(\theta)\, dY, \\ S(\theta_0) &= 0,\ S(\theta_1) = E. \end{aligned} \tag{32}$$

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

$$\max \int_{\theta_0}^{\theta_1} \int_{Y \in \Gamma} h(Y) \left[u(Y - \tau - f) + y + \tau + f - ke\right] \phi(\theta)\, dY d\theta$$

$$s.t. \quad \dot{S}(\theta) = \int_{Y \in \Gamma} he\phi dY, \quad U_\theta = y_\theta - \psi_\theta m,$$

$$m_\theta = x, \quad e_\theta = z, \quad U = u(Y - \tau - f) + y - \psi m + t,$$

$$y_{e\theta} z - \psi_\theta x \geq 0, \quad U \geq P(\theta, Y - \tau, E), \quad m \geq 0, \quad t \geq 0,$$

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

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

From the first order conditions, it follows that:

$$\rho^1(\theta,Y)=h(Y)\,\phi(\theta)\left[1-\frac{1}{u'(Y-\tau-f)}\right], \tag{33}$$

$$\lambda^2(\theta,Y)=0, \text{ for almost all } Y,\theta, \tag{34}$$

$$\mu^3(\theta,Y)=\lambda^1(\theta,Y)\,\psi_\theta, \tag{35}$$

$$\rho^1(\theta,Y)=\lambda^4(\theta,Y), \tag{36}$$

$$\mu^4(\theta,Y)=-\lambda^1(\theta,Y)\,y_{e\theta}, \tag{37}$$

$$\rho^1(\theta,Y)=-\mu^2_\theta(\theta,Y), \tag{38}$$

$$
\begin{aligned}
-\mu^4_\theta(\theta,Y) \quad = \quad & h(Y)\,\phi(\theta)\left[y_e-k+\mu^1(\theta,Y)\right]+\mu^2(\theta,Y)\ y_{e\theta} \\
& -\rho^1(\theta,Y)\ y_e+\lambda^1(\theta,Y)\ z\ y_{ee\theta},
\end{aligned} \tag{39}
$$

$$\mu^1(\theta)=\mu^1 \text{ constant}, \tag{40}$$

$$-\mu^3_\theta(\theta,Y)=-\mu^2(\theta,Y)\,\psi_\theta+\rho^1(\theta,Y)\,\psi+\lambda^3(\theta,Y). \tag{41}$$

$$\min\{\lambda^2;U-P(\theta,Y-\tau,E)\}=0 \tag{42}$$

Substituting (33) in (38),

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

Notice that equation (34) implies that $U(\theta,Y)\geq P(\theta,Y-\tau(Y),E)$ is never binding. Hence, as $U(\theta_0)$ and $U(\theta_1)$ are free, the transversality condition implies that

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

Then, as $\lambda^4(\theta, Y) \geq 0$, from equation (36) we get $u'(Y - \tau - f) \geq 1$, for almost all Y, θ.

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

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

Therefore, the second-best amount of consumption is the same as the first-best amount and one must (almost) always face a unitary marginal tax. Moreover, from equation (38),

$$\mu^2(\theta_1, Y) = \rho^1(\theta, Y) = 0, \text{ for almost all } Y, \theta. \tag{43}$$

Then, from equation (41), we get

$$\mu^3_\theta(\theta, Y) = -\lambda^3(\theta, Y) \leq 0. \tag{44}$$

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

$$\mu^3(\theta, Y) = \lambda^3(\theta, Y) = 0, \text{ for almost all } Y, \theta. \tag{45}$$

Therefore, equations (35) and (3) imply $\lambda^1(\theta, Y) = \mu^4(\theta, Y) = 0$. Then, from equation (39) we get

$$\mu^1 = k - y_e. \tag{46}$$

Lemma 24 *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 12. As $\frac{\partial W}{\partial E}\Big|_{e=e^{2b}, E=E^{2b}} = \mu^1(\theta_1) = \mu^1$, it follows that

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

Substituting into (46),

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

which is the equation that defines e^*. ■

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

We have to show that the monotonicity condition (8) is satisfied in the relaxed problem considered. But, as we have already shown in remark (7), $e^*(\theta)$ is increasing in θ. Therefore the monotonicity condition $e_\theta(\theta, Y) \geq 0$ is satisfied.

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

Proposition 12 follows as a corollary since it is a particular case of the economy with default when $\psi(\theta, Y) = 1$ for all θ, Y.

2.1.3 Existence and uniqueness of equilibria:

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

Proposition 25 *There exists e^u and e^c such that $k = y_e(\theta, e^u; E)$, $ku'(Y - ke^c) = y_e(\theta, e^c; E)$. Moreover, e^u and e^c are unique.*

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

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

The same argument establishes the existence of the education profile defined before Lemma 17 and in equation (27).

Corollary 26 *There exists a unique e^P such that $\hat{k} = y_e\left(\theta, e^P(\theta, Y); E\right)$.*

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

Proposition 27 *There exists a unique e^* such that*

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

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

Define the function ρ as

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

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

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

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

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

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

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

References

ARAÚJO, A. and MOREIRA, H. (2000). "Adverse Selection Problems without the Spence-Mirrlees Condition" mimeograph, Getulio Vargas Foundation.

ARAÚJO, A. and MOREIRA, H. (2001). "Non-monotone insurance contracts and their empirical consequences" mimeograph, Getulio Vargas Foundation.

ARAÚJO, A., GOTTLIEB, D., and MOREIRA, H. (2004a). "The Empirical Content of the Signaling Hypothesis" mimeograph, Getulio Vargas Foundation.

ARAÚJO, A., GOTTLIEB, D., and MOREIRA, H. (2004b). "A model of mixed signals with applications to countersignaling and the GED" mimeograph, Getulio Vargas Foundation.

ARAÚJO, A., MOREIRA, H., and TSUCHIDA, M. (2003). "Do dividends signal more earnings? A theoretical analysis" mimeograph, Getulio Vargas Foundation.

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

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

BARRICK, M. and MOUNT, M. (1991). "The Big Five personality dimensions and job performance: a meta-analysis", *Personnel Psychology*, **44**, 1-26.

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

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

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

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

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

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

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

BOUDREAU, J., BOSWELL, W., and JUDGE, T. (2001). "Effects of personality on executive career success in the United States and Europe", *Journal of Vocational Behavior*, **58**, 53-81.

BOWLES, S. and GINTIS, H. (1976). *Schooling in Capitalist America* (New York: Basic Books).

BOWLES, S. and GINTIS, H. (1998) "The Determinants of Individual Earnings: Cognitive Skills, Personality and Schooling", unpublished manuscript, University of Massachussets, Amherst.

BOWLES, S. and GINTIS, H. (2001) **,** "Schooling in Capitalist America Revisited", *Sociology of Education*, **75**, 1-18.

BURTLESS, G. (1985). "Are Targeted Wage Subsidies Harmful? Evidence from a Wage Voucher Experiment", Industrial and Labor Relations Review, **39**, 105-114.

CAMERON, S. and HECKMAN, J. (1993). "The Nonequivalence of High School Equivalents", *Journal of Labor Economics*, January, **11**, 1-47.

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

CARNEIRO, P. and HECKMAN, J. (2003). "Human Capital Policy", Woking Paper 9495, National Bureau of Economic Research.

CAWLEY, J., CONNEELY, K., HECKMAN, J., and VYTLACIL; E. (1996). "Measuring the Effects of Cognitive Ability", Woking Paper 5645, National Bureau of Economic Research.

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

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

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

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

DEWATRIPONT, M., JEWITT., I., and TIROLE, J. (1999). "The Economics of Career Concerns, Part I: Comparing Information Structures", *Review of Economic Studies*, **66**, 183-198.

DRAZEN, A. and HUBRICH, S. (2003). "Mixed Signals in Defending the Exchange Rate: What do the Data Say?", CEPR Discussion Papers, #4050.

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

DUNIFON, R. and DUNCAN, G. (1996). "Long-Run Effects of Motivation on Labor-Market Success", *Social Psychology Quarterly*, 1996, **61**, 33-48.

DUNIFON, R. and DUNCAN, G. (1999). "Soft Skills and Long-Run Market Success", *Research in Labor Economics*, 1-42.

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

EDWARDS, R. (1976) "Individual Traits and Organizational Incentives: What Makes a Good Worker?", *Journal of Human Resourses*, **11**, 51-68.

ENGERS, M. (1987). "Signalling with Many Signals", *Econometrica*, **55**, 663-674.

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

FELTOVICH, N., HARBAUGH, R., and TO, T. (2002). "Too Cool for School? Signaling and Countersignaling", *Rand Journal of Economics*, **33**, 630-649.

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

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

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

FLEURBAEY, M., GARY-BOBO, R., and MAGUAIN, D. (2002) "Education, distributive justice, and adverse selection", *Journal of Public Economics*, **84**, 113-150.

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

GARDNER, H. (1993). *Multiple Intelligences: The Theory in Practice* (Basic Books, New York).

GOFFIN, R., ROTHSTEIN, M., and JOHNSTON, N. (1996). "Personality testing and the assessment center: Incremental validity for managerial selection", *Journal of Applied Psychology*, **81**, 746-756.

GOTTLIEB, D., and MAESTRI, L. (2004). "Banning Information as a Redistributive Device" mimeograph, Getulio Vargas Foundation.

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

HECKMAN, J. and RUBINSTEIN, Y. (2001). "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *AEA Papers and Proceedings*, **91**, 145-9.

HECKMAN, J., HSEE, J., and RUBINSTEIN, Y. (2000). "The GED is a Mixed Signal", Unpublished manuscript presented at American Economic Association meeting, Boston, Massachusetts.

HOGAN, J. and HOGAN, R. (1989) "How to measure employee reliability", *Journal of Applied psychology*, 1989, **74**, 73 - 79.

HOLMSTROM, B. (1999). "Managerial Incentive Problems: A Dynamic Perspective", *Review of Economic Studies*, **66**, 169-182.

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

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

KLEIN, R., SPADY, R., and WEISS, A. (1991). "Factors Affecting The Output and Quit Propensities of Production Workers", *Review of Economic Studies*, **58**, 929-954.

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

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

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

KOHLLEPEL, L. (1983). "Multidimensional 'Market Signaling"' Discussion Paper, Universität Bonn.

LAURENCE, J. (1998) "Use of The GED by the United States Armed Forces", in: *Heckman, James*, editor, The GED, forthcoming.

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

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

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

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

MURNANE, R., WILLET, J., and TYLER, J. (2000). "What Are the High School Diploma and the GED Worth in the Labor Market? Evidence for Males from High School and Beyond", *Review of Economics and Statistics,* **82**, 23-37.

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

O'NEAL, B. (2002). "Nuclear Weapons and the Pursuit of Prestige", mimeograph, UCLA.

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

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

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

QUINZII, M. and ROCHET, J. (1985). "Multidimensional Signaling", *Journal of Mathematical Economics*, **14**, 261-284.

SPENCE, M. (1973). *Market signalling: Information transfer in hiring and related processes* (Harvard University Press, Cambridge, MA).

SPENCE, M. (1974). ."Competitive and optimal responses to signals: An Analysis of Efficiency and Distribution", *Journal of Economic Theory*, **7**, 296-332.

STERNBERG, R. (1985). *Beyond IQ: A Triarchic Theory of Human Intelligence*, (Cambridge University Press: Cambridge).

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

RILEY, J. (1979). "Informational equilibrium" *Econometrica*, **47**, 331-359.

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

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

TYLER, J., MURNANE, R., and WILLET J. (2000). "Estimating the Labor Market Signaling Value of the GED", *Quarterly Journal of Economics* **115**, 431-468.

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

WEISS, A. (1988). "High School Graduation, Performance and Wages" *Journal of Political Economy,* **96**, 785-820.