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

Aloisio Araujo† Daniel Gottlieb‡ Humberto Moreira§

June 25, 2004

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 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

*The usual disclaimers apply. Comments are welcome.

†IMPA and Graduate School of Economics, Getulio Vargas Foundation.

‡Graduate School of Economics, Getulio Vargas Foundation. Corresponding Author: gottlieb@fgvmail.br.

§Graduate School of Economics, Getulio Vargas Foundation.

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 (2002) 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 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), 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 Ro-

[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.

chet 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 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. Section 2 presents the basic framework. Section 3 analyzes the pure signaling case and Section 4 characterizes the equilibrium. Section 5 discusses how countersignaling may emerge.

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

Section 6 employs this framework to understand the GED exam. Then, Section 7 concludes. Appendix B extends the results to the case where schooling enhances productivity.

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 $p : \Theta \to \mathbb{R}_{++}$.

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 $f(\iota, \eta, y)$ and act competitively. Moreover, other than schooling, firms have an interview technology which is a non-sufficient statistic of the worker's type $g(\iota, \eta)$. Thus, even though firms have some idea of the overall ability of a worker, they are unable to unambiguously determine her characteristics (i.e., g is not a one-one function).[4]

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:*

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

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

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

[4] 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).

Next, we will specify the analytical forms of the functions presented.[5] 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.[6]

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.

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

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

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.[7] 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.

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

[6] In the Appendix D, we consider the case where schooling distorts the worker's performance in the interview.

[7] 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).

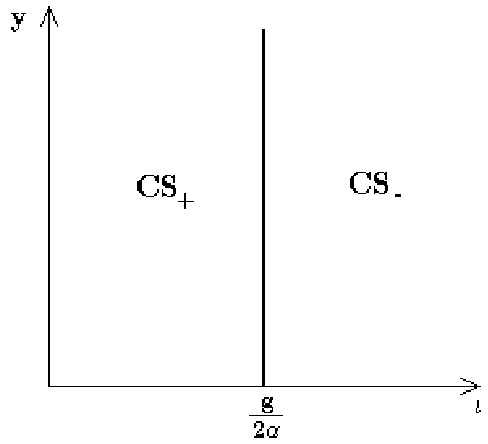

Figure 1

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

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

where $a, b \in \mathbb{R}_+$. If $b > 1 - a - b$ we say that the firm's technology is intensive in cognitive skills. Otherwise, we say that it is intensive in non-cognitive skills.

When $a = 0$, the model is called a pure signaling model as the workers' productivity is not affected by schooling.

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

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

Throughout most of the paper, we will focus on the pure signaling case ($a = 0$). Thus, we will omit the dependence of s on y.[8]

3 The pure signaling case

In this section, we follow Spence (1973) and assume that education is a costly signal that does not enhance the worker's productivity. In the next subsection, we will characterize the signaling equilibria. Then, we will 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.

3.1 The signaling equilibria

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).*

[8] We study the productive-schooling case ($a > 0$) in the Appendix B.

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$-$ (ι, 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$-$ (ι, 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$-$ (ι, 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.

3.1.1 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

$$y_\iota(\iota, g) = s(\iota)(bg - \alpha\iota),$$

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:

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

$$\begin{aligned} y_\iota(\iota, g) &= s(\iota)(bg - \alpha\iota), \\ y_\iota(\iota, g)(g - 2\alpha\iota) &\geq 0. \end{aligned} \tag{7}$$

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 the first equation of 7, it follows that

$$y_\iota(\iota, g) > 0 \iff \iota < \frac{bg}{\alpha} = \iota^*(g). \tag{8}$$

■

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

Remark 1 *Notice that the second equation of (7) and equation (6) imply that*

$$y_\iota(\iota, g) \geq 0 \Longleftrightarrow \iota \leq \frac{g}{2\alpha}. \tag{9}$$

Generally, equations (8) and (9) 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).*

3.1.2 Continuous pooling set

Henceforth, we will assume that, conditional of the result of the interview g, intelligence is uniformly distributed. Although this assumption is not necessary for any of our results, it greatly simplifies the algebra.

Assumption 1 $\iota \mid g \sim U[\iota_0, \iota_1]$.

Suppose there exists a non-degenerate interval $I = [\iota_a, \iota_b]$ which is a continuous pooling set and that no closed interval $X \supset I$ is a continuous pooling set. Then, $y(\iota, g) = \bar{y}(g)$ for all $\iota \in I$.

The zero-profit condition is

$$w(\bar{y}(g), g) = \frac{1}{\iota_b - \iota_a} W(\iota_b, \iota_a, g), \tag{10}$$

where $W(\iota, \iota_a, g) = \int_{\iota_a}^{\iota} f(x, g - \alpha x)\, dx$. Conditions 2 and 3 from Definition 1 are thrivially satisfied in that given interval.

3.1.3 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 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 workers with types ι and $\tilde{\iota}$ choose the same level of education and have the same interview result, then their marginal cost of education must be the same:*

$$y(\iota, g) = y(\tilde{\iota}, g) \Rightarrow \frac{\partial c(\iota, g, y)}{\partial y} = \frac{\partial c(\tilde{\iota}, g, y)}{\partial y}.$$

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:

$$\hat{\iota} = \frac{g}{\alpha} - \iota \equiv \gamma(\iota, g). \tag{11}$$

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 ($b > \frac{1}{2}$).*

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

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, it satisfies:*

$$y_\iota(\iota, g) = \frac{1}{2}\left\{f(\iota, \eta)(bg - \alpha\iota) + \alpha^{1-2b} f(\eta, \iota)\left[(1-b)g - \alpha\iota\right]\right\}, \tag{12}$$
$$y_\iota(\iota, g)(g - 2\alpha\iota) \geq 0.$$

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

3.2 Equilibrium selection and comparative statics

The proposition below presents some comparative statics results. As overall ability provides a higher productivity, the zero-profit condition implies that wages must be strictly increasing in g. Moreover, since education is costly, individuals would only choose to educate if this increases their wages.

Proposition 1 *Wages are strictly increasing*

1. *in the result of the interview g, and*
2. *in the amount of schooling (controlling for the interview g).*

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 in practice. 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.*

4 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 11, we denote by $\gamma(\iota, g)$ the type with the same marginal cost of signaling as ι. We will focus on the case where $\gamma(\iota_0, g) \leq \iota_1$ and $b < 1/2$ (the other cases can be studied in a similar fashion).[10] Clearly, as $\gamma(\iota_0, g) \leq \iota_1$,

[10] See Araujo, Gottlieb, and Moreira (2004) for a characterization of more general models.

it follows that $(\gamma(\iota_0, g), \iota_1]$ must be a separating set in any quasi-separable equilibrium (as no other type can have the same marginal cost of schooling as $\iota \in (\gamma(\iota_0, g), \iota_1]$). The characterization will be done through a series of lemmata.

Define $\gamma(\iota) \equiv \gamma(\iota, g)$. The first lemma provides another necessary condition for truth-telling. This additional condition will determine the boundary condition for the amount of education when changing from discrete pooling sets to separating sets.

Lemma 5 *(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:*

$$y(\iota) = -\frac{\iota(g-\alpha\iota)[s(\iota) - s(\gamma(\iota))]}{2} + \lim_{x \to \iota_-} y(x). \tag{13}$$

(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:

$$y(\iota) = -\frac{\iota(g-\alpha\iota)[s(\iota) - s(\gamma(\iota))]}{2} + \lim_{x \to \iota_+} y(x). \tag{14}$$

Lemma 5 can be restated as claiming that a discontinuous indirect utility function cannot be truth-telling. The basic intuition behind this result is that, as the cost of signaling is continuous, if 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 second lemma determines the discrete pooling set.

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

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

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, $y(\iota_1, g) = 0$.*

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 ensures that the necessary conditions are also sufficient for the equilibrium.

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, 5, 7, and 8.

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* $\iota > \gamma_0$;
2. $y(\iota_1) = 0$;
3. $y_\iota(\iota, g) = \frac{1}{2}\left\{f(\iota, \eta)(bg - \alpha\iota) + \alpha^{1-2b} f(\eta, \iota)\left[(1-b)g - \alpha\iota\right]\right\}$, *for* $\iota \leq \gamma_0$;
4. $y(\gamma_0) = \frac{\iota_0(g - \alpha\iota_0)[s(\iota_0) - s(\gamma(\iota_0))]}{2} + \lim_{x \to \gamma(\iota_0)+} y(x)$.

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 functions, 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 γ_0 since $s(\iota_0) - s(\gamma_0) > 0$ (see Lemma 3 and $b < 1/2$). This follows from the fact that wages are discontinuous: individuals with $\iota \in \left[\frac{g}{2\alpha}, \gamma_0\right]$ earn wages higher than their productivity since they are pooled with more productive workers but those with types higher than γ_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 non-implementable. 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 $b = 0.4$, $g = 10$, $\alpha = 1$, $\iota_0 = 1$, $\iota_1 = 10$.[11]

[11] The equilibrium profiles of education, wages, and utility for other parameters are represented in the Appendix 1.E.

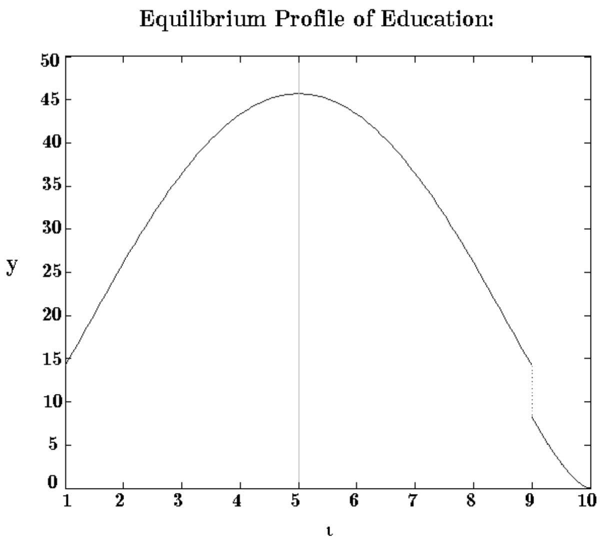

Figure 2

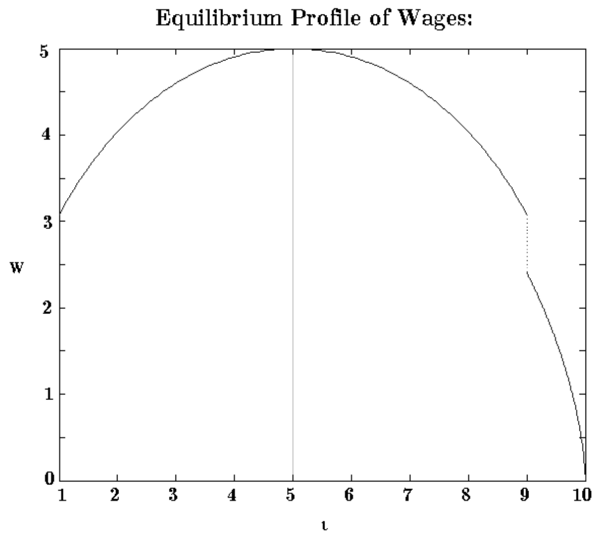

Figure 3

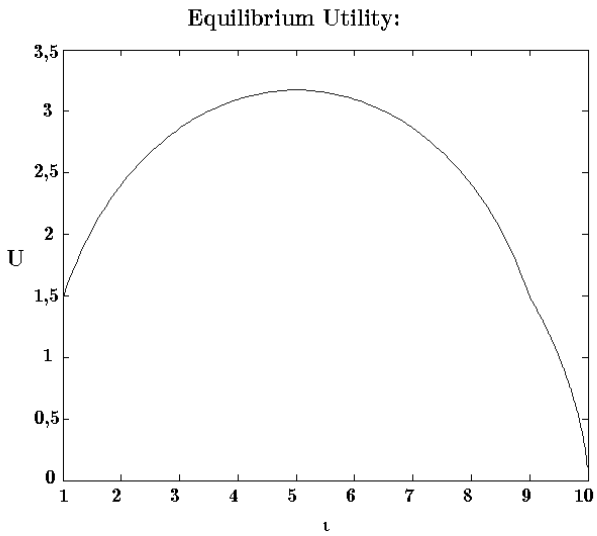

Figure 4

Notice that both education and wages are discontinuous but the utility is countinous in ι.

5 Countersignaling

In some situations, high-productivity individuals may choose to signal less than those with lower productivity. Clements (2002) 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, pp.18) 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, 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 of 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 4, 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 schooling 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 SCP does not hold (the schooling and the firms' technologies are not the same), 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.

6 The GED exam

6.1 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)].

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.

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.[12]

6.2 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

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

[12] 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)].

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 substitution between intelligence and perseverance.[13] We assume that there is no cost in taking the test.[14]

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

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

The following assumption states that the exam emphasizes intelligence more than the interview g does.[15]

Assumption 2 *$\kappa > \alpha$.*

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 $\frac{\overline{g}-g}{\kappa-\alpha} < \frac{g}{2\alpha}$.

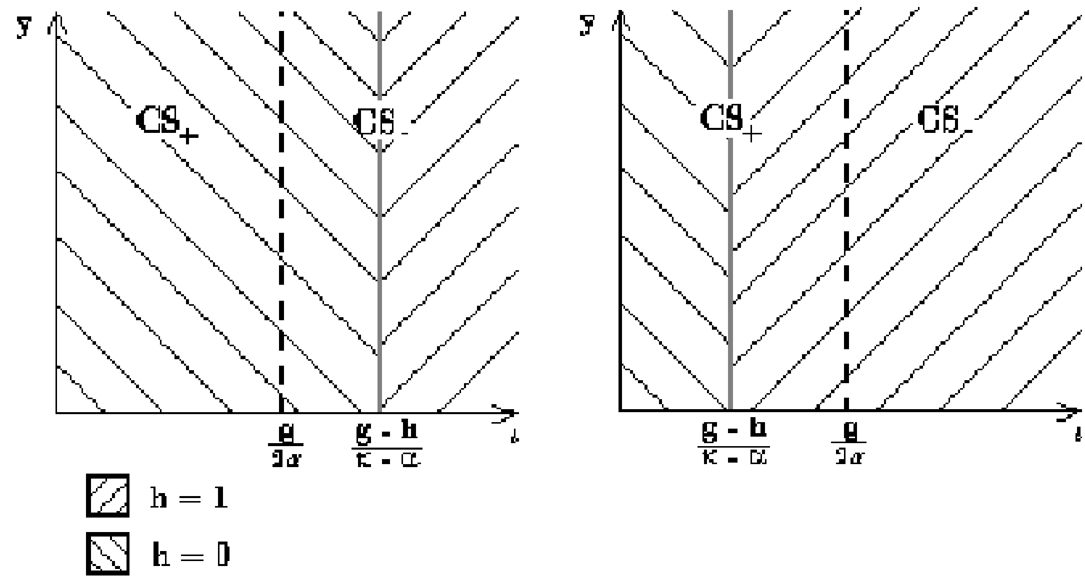

Figure 5

[13] 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.

[14] 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.

[15] According to Heckman, Hsee and Rubinstein (1993), this is the case as the GED exam is intensive in cognitive skills.

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.

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 $\iota \leq \frac{\overline{g}-g}{\kappa-\alpha} \leq \hat{\iota}$ are pooled in the same contract (otherwise, the signal is not informational). Then type-$\hat{\iota}$ 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-$\hat{\iota}$ worker's strategy is not optimal (condition 1).

Thus, any wage schedule such that a type-$\hat{\iota}$ individual earns less than type ι cannot be an equilibrium. Hence, a type-$\hat{\iota}$ 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 the GED 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.[16] However, as we show in the

[16] 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.

next proposition, when the technology is intensive in cognitive skills, the GED signal may be non-neutral in equilibrium.

Proposition 6 *If the firms' technology is intensive in cognitive skills and there are two types pooled in the same contract such that* $\iota \leq \frac{\overline{g}-g}{\kappa-\alpha} \leq \hat{\iota}$, *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 2 *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)].*

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 C. ■

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 the design of optimal tests.

7 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 on 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.

Define ψ as
$$\psi(\iota, \eta, \bar{g}) = g(\iota, \eta) - \bar{g}.$$

From the inverse function theorem, there exists φ such that
$$\varphi(\iota, \bar{g}) = \eta.$$

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

Substituting in the cost function, it follows that
$$c = \frac{y}{w(\iota, \varphi(\iota, \bar{g}))}.$$

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

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:[17]
$$\begin{aligned} w &= \left[\alpha_1 \iota^\rho + \alpha_2 \eta^\rho\right]^{\frac{1}{\rho}}, \\ g &= \left[\beta_1 \iota^\gamma + \beta_2 \eta^\gamma\right]^{\frac{1}{\gamma}}. \end{aligned}$$

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.

[17] 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 The productive schooling case

In this section, we extend the results from the pure signaling case to the productive schooling case ($a > 0$). The main conclusion is that all the basic results still hold when schooling is productive.

B.1 The signaling equilibria

As in Section 3.1, we will obtain necessary conditions for an equilibrium in separating sets, continuous pooling sets and discrete pooling sets separately.

B.1.1 Separating set

The zero-profit and consistency of beliefs conditions imply

$$w\left(y\left(\iota,g\right),g\right)=f\left(\iota,g-\alpha\iota,y\left(\iota,g\right)\right).$$

The truth-telling, the following condition is

$$\iota\in\arg\max_{\hat{\iota}}f\left(\tilde{\iota},g-\alpha\tilde{\iota},y\left(\tilde{\iota}\right)\right)-c\left(\iota,y\left(\hat{\iota}\right)\right).$$

As in Lemma 1, the following lemma follows from the local first- and second-order conditions of the problem above.

Lemma 9 *In any separating set, if a C^2 by parts education and wage profile is truth-telling it must satisfy:*

$$y_{\iota}\left(\iota,g\right)=y_{\iota}\left(\iota,g\right)=s\left(\iota,y\left(\iota\right)\right)y\left(\iota,g\right)\frac{bg-\left(1-a\right)a\iota}{y\left(\iota,g\right)-as\left(\iota,y\left(\iota\right)\right)},\qquad(15)$$
$$y_{\iota}\left(\iota,g\right)\left(g-2\alpha\iota\right)\geq0.$$

B.1.2 Continuous pooling set

As before, we will assume that conditional on g, intelligence is uniformly distributed (Assumption 1).

Suppose there exists a non-degenerate interval $I=[\iota_a,\iota_b]$ which is a continuous pooling set and that no closed interval $X\supset I$ is a continuous pooling set. Then, $y\left(\iota,g\right)=\bar{y}\left(g\right)$ for all $\left(\iota,g\right)\in I$. Thus, the zero-profit condition is

$$w\left(\bar{y}\left(g\right),g\right)=\frac{1}{\iota_a-\iota_b}\tilde{W}\left(\iota_b,\iota_a,g,\bar{y}\left(g\right)\right),$$

where $\tilde{W}\left(\iota,\iota_a,g,y\right)=\int_{\iota_a}^{\iota}f\left(x,g-\alpha x,y\right)dx$

Conditions 2 and 3 are thrivially satisfied in that given interval.

B.1.3 Discrete pooling set

As argued in the pure signaling case, a necessary condition for truth-telling in a discrete pooling set is the marginal rate of substitution identity (Lemma 2). Thus, if two workers (ι, g) and $(\hat{\iota}, g)$ choose the same level of education, it follows that

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

The zero-profit and consistency of beliefs conditions imply that

$$w(y(\iota, g), g) = \left[s(\iota, g, y(\iota)) + \frac{\alpha^{1-2b}\iota(g - \alpha\iota)}{s(\iota, y(\iota, g))} \right].$$

The truth-telling condition is

$$\iota \in \arg\max_{\hat{\iota}} \frac{1}{2} \left[s(\tilde{\iota}, y(\tilde{\iota}, g)) + \alpha^{1-a-2b} y^a (g - \alpha\tilde{\iota})^b \tilde{\iota}^{1-a-b} \right] - c(\iota, y(\hat{\iota}, g))$$

Then, the next lemma is similar to Lemma 4.

Lemma 10 *If (ι, g) belongs to a discrete pooling set, then if a C^2 by parts education and wage profile is truth-telling, it satisfies:*

$$y_\iota(\iota, g) = y(\iota, g) \times \frac{f(\iota, \eta, y)[b\eta - (1-a-b)\alpha\iota] + f(\eta, \iota, y)\alpha^{1-a-2b}[(1-a-b)\eta - bg]}{y(\iota, g) - \alpha\iota\eta[f(\iota, \eta, y) + \alpha^{1-a-2b} f(\eta, \iota, y)]},$$

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

The following lemma is analogous to Lemma 3.

Lemma 11 *If two workers are discretely pooled, then the less intelligent one is more productive if the firms' technology is intensive in perseverance ($b < 1-a-b$) and the more intelligent one is more productive if the firms' technology is intensive in intelligence ($b > 1-a-b$).*

Proof. 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,

$$f(\iota, g, y) > f(\hat{\iota}, g, y) \iff y^a \iota^b \hat{\iota}^{1-a-b} > y^a \hat{\iota}^b \iota^{1-a-b} \iff b > 1-a-b.$$

■

B.1.4 Equilibrium selection

As in the pure signaling case, we apply the quasi-separability criterion (see Definition 4). The following proposition states that the result of Proposition 2 can be extended to the productive schooling case and can be seen as an evidence of the failure of the SCP.

Proposition 8 *If the pooling set is non-empty, then the SCP does not hold.*

Proof. Identical to the proof of Proposition 2. ■

B.1.5 Countersignaling

It is straightforward to show that, just like in the pure signaling case, education is strictly increasing for $\iota < \frac{g}{2\alpha}$ and strictly decreasing for $\iota > \frac{g}{2\alpha}$. Moreover, as in page 8, it can be shown that

$$\frac{bg}{\alpha(1-a)} = \arg\max \arg\max y^a \iota^b \eta^{1-a-b} \quad \text{s.t. } g = \alpha\iota + \eta.$$

Hence, the productivity of a worker with interview result g is strictly increasing for $\iota < \frac{bg}{\alpha(1-a)}$ and strictly decreasing for $\iota > \frac{bg}{\alpha(1-a)}$. Then, the countersignaling interval is $[\frac{bg}{\alpha(1-a)}, \frac{g}{2\alpha}]$ if $b < \frac{1-a}{2}$ or $[\frac{g}{2\alpha}, \frac{bg}{\alpha(1-a)}]$ if $b > \frac{1-a}{2}$. Hence, countersignaling does not occurs if, and only if, $b = \frac{1-a}{2}$.

B.2 The GED exam

In this subsection, we analyze the effect of introducing the GED exam in the productive schooling model. As before, the GED consists of a test where individuals with a sufficiently high combination of characteristics pass.

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

where $\kappa > \alpha$ (Assumption 2) and $\overline{g} \in \mathbb{R}_{++}$.

The next proposition generalizes Propositions 5 and 6 to the productive schooling case.

Proposition 9 *If the firms' technology is intensive in non-cognitive skills ($b < 1 - a - b$), the introduction of the GED exam does not modify the wage schedule.*

If the firms' technology is intensive in cognitive skills ($b > 1 - a - b$) and there are two types pooled in the same contract such that $\iota \leq \frac{\overline{g}-g}{\kappa-\alpha} \leq \hat{\iota}$, 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. Analogous to the proof of Propositions 5 and 6. ■

Thus, even though the GED conveys information about workers' characteristics, it does not affect the equilibrium wage schedule. The main problem of the GED is its emphasis on cognitive skills. If it were intensive in non-cognitive skills ($\kappa < \alpha$) it would be a non-neutral signal since the individuals able to pass would be the more productive ones.

C Number of tests required for full-separability

As shown in Section 6, 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 6 as well as Engers and Fernandez's (1987) result that when the single-crossing property holds no additional signal is required.

The following assumptions 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.

Assumption A.2 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 ensures the existence of equilibrium.

Assumption A.3 $p, f \in C^1$.

The following proposition states that under Assumptions A.1, A.2, and A.3, there always exists a quasi-separable equilibrium.

Proposition 10 *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 $\eta \leq n+1$ be the number of pooled types.

Introduce a costless signal h_1 that allows the type with the lowest productivity to obtain $h_1 = 1$ and that assigns $h_1 = 0$ to all other types (if more than one type have the same productivity, take any of them). Then, the new equilibrium will only feature $\eta - 1$ pooled types and incentive-compatibility implies that the profile of equilibrium wages remain unchanged.

Repeating the process t times, there will be at most $\eta - t$ pooled types. Thus, introducing $\eta - 1$ new signals, it follows that there will be at most 1 type pooling in each contract. ■

D Case where schooling affects the interview

In this section, we consider the case where the overall ability measure also reflects an individual's schooling. This can be the case, for example, if schooling distorts the ability perceived by the firms. Hence, we define $g(\iota, \eta, y)$ as

$$g(\iota, \eta, y) = \alpha\iota + \beta y + \eta,$$

where α and β are the marginal rates of substitution between non-cognitive skills and cognitive skills and schooling.

Substituting into the cost function, we get

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

$$
\begin{aligned}
c_y &= \frac{1}{\iota\left(g-\alpha\iota-\beta y\right)}+\frac{\beta y}{\iota\left(g-\alpha\iota-\beta y\right)^2}. \\
c_{y\iota} &= -\frac{1}{\left(\iota\eta\right)^2}\left(g-2\alpha\iota-\beta y\right)-\frac{\beta y}{\iota^2\eta^3}\left(g-3\alpha\iota-\beta y\right) \\
&= -\frac{1}{\iota^2\eta^3}\left[\left(g-\alpha\iota-\beta y\right)\left(g-2\alpha\iota-\beta y\right)+\beta y\left(g-3\alpha\iota-\beta y\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
c_{y\iota}=0 &\Longleftrightarrow \left(g-\alpha\iota-\beta y\right)\left(g-2\alpha\iota-\beta y\right)+\beta y\left(g-3\alpha\iota-\beta y\right)=0 \\
&\Longleftrightarrow 2\alpha\iota^2-3g\iota+\frac{g}{\alpha}\left(g-\beta y\right)=0 \\
&\Longleftrightarrow 2\alpha^2\iota^2-3\alpha\iota g+g^2=\beta g y
\end{aligned}
$$

Solving the equation above, it follows that

$$
\iota=\frac{3g\pm\sqrt{g\left(g+8\beta y\right)}}{4\alpha},
$$

$$
\left.\frac{dy}{d\iota}\right|_{c_{y\iota}=0}=\frac{4\alpha^2}{\beta g}\left(\iota-\frac{3g}{4\alpha}\right).
$$

Hence, the parameter space can be divided into 2 regions according to the signal of $c_{y\iota}$: CS_+ and CS_-.

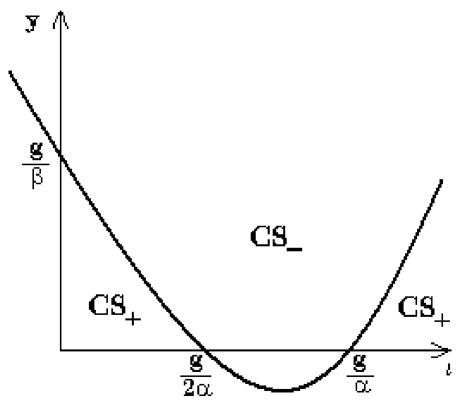

Figure 6

As usual, a necessary condition for truth-telling is that the signal must be increasing in the CS_+ region and decreasing in the CS_- region.

If two individuals are discretely pooled, they must have the same marginal cost of signaling:

$$c_y(\iota, \eta, y) = c_y(\tilde{\iota}, \tilde{\eta}, y).$$

This condition yields

$$(\eta + \beta y)\tilde{\eta}^2 + \left[\eta^2 + (2\beta y - g)\eta + (\beta y - g)\beta y\right]\tilde{\eta} + \beta y \eta\left[\eta + \beta y - g\right] = 0. \quad (16)$$

If equation (16) has two different roots ($\tilde{\eta}_1 < \tilde{\eta}_2$), then $\tilde{\eta}_1$, $\tilde{\eta}_2$ and η may be discretely pooled. If it has one root, η may discretely pool with at most one different type in equilibrium. Otherwise, η cannot be discretely pooled.

The following proposition establishes that the GED signal is neutral if the relative intensity of non-cognitive skills is high.

Proposition 11 *There exists $\bar{b} > 0$ such that, for every $b \leq \bar{b}$, the introduction of the GED exam does not modify the wage schedule.*

Proof. The result is trivial for a separating set. Assume two workers with types $\iota < \hat{\iota}$ are pooled in the same contract and that $\iota \leq \frac{\overline{g}-g}{\kappa-\alpha} \leq \hat{\iota}$ (otherwise, the signal is not informational). Hence, type-$\hat{\iota}$ has $h = 1$ (if he chooses to take the exam) and type-ι has $h = 0$. But, from the definition of g, it follows that

$$\eta = g - \alpha\iota - \beta y > g - \alpha\hat{\iota} - \beta y = \hat{\eta}.$$

Thus, there exists $\bar{b} \in [0, 1]$ such that $y^a \iota^b \eta^{1-a-b} < y^a \hat{\iota}^b \hat{\eta}^{1-a-b}$, for all $b \geq \bar{b}$.

But, as a wage schedule such that a type-$\hat{\iota}$ individual earns less than type ι is not truth-telling, a type-$\hat{\iota}$ individual must earn the same as type ι and is indifferent between taking the GED or not. ■

Remark 8 *If the firms' technology is sufficiently intensive in cognitive skills, then the neutrality of the GED does not hold. The economic intuition for this result is that, if passing the test signals higher productivity, then individuals can use it to truthfully reveal their types.*

E Graphs of the Equilibrium with different parameters

Figures 8, 9 and 10 show the equilibrium amounts of education, wages, and utility for $b = \frac{1}{4}$, $g = 10$, $\alpha = 1$, $\iota_0 = 2$, $\iota_1 = 9$.

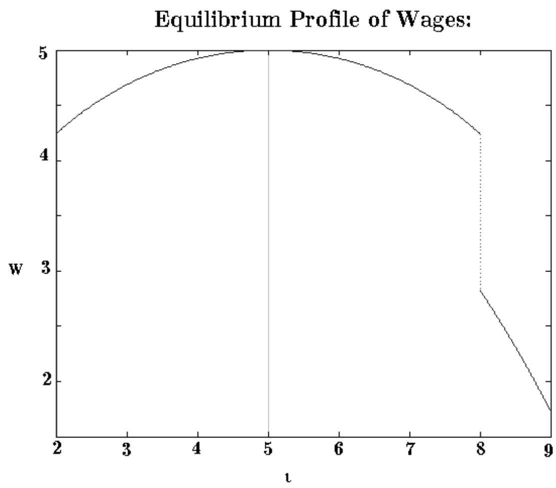

Figure 1: Figure 8

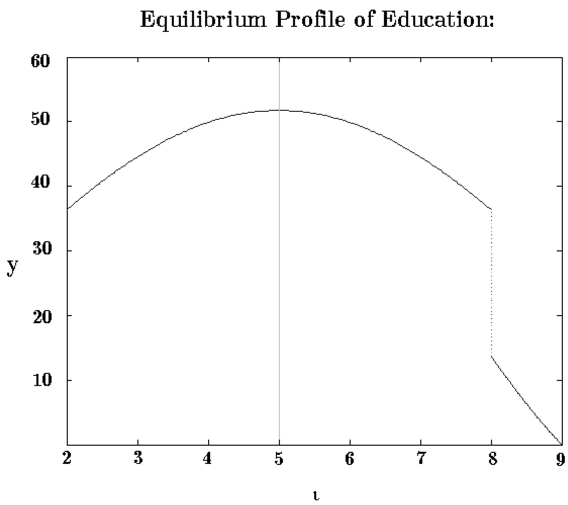

Figure 7

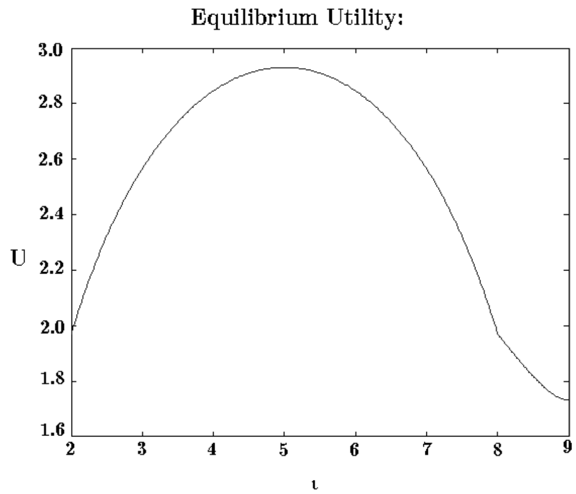

Figure 9

Figures 11, 12 and 13 show the equilibrium amounts of education, wages, and utility for $b = \frac{1}{4}$, $g = 8$, $\alpha = \frac{4}{5}$, $\iota_0 = 2$, $\iota_1 = 9$.

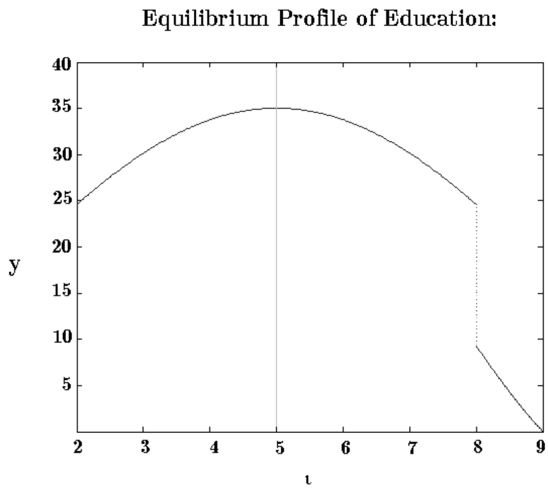

Figure 10

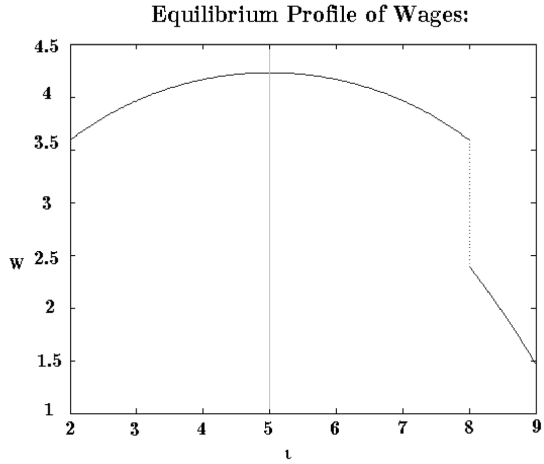

Figure 11

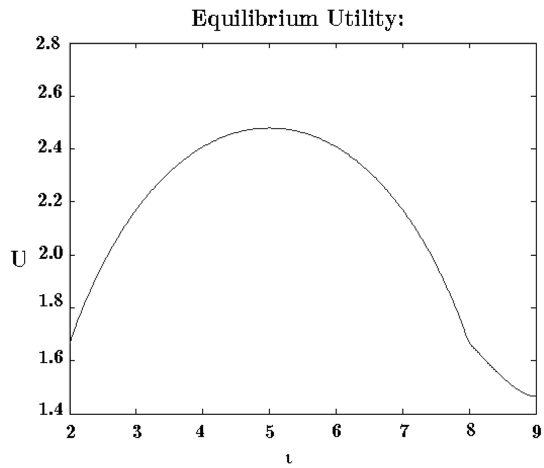

Figure 12

F Proofs

Proof of Lemma 1:

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

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

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

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

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

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

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

The first-order condition yields, for all ι,

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

Differentiating the condition above, we get

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

$$- c_y(\iota, g, y(\iota, g))\, y_{\iota\iota}(\iota, g). \tag{20}$$

The second-order condition yields

$$s_{\iota\iota}(\iota) - c_{yy}(\iota, g, y(\iota, g))\, y_\iota(\iota, g) - c_y(\iota, g, y(\iota, g))\, y_{\iota\iota}(\iota, g) \leq 0. \tag{21}$$

Substituting (19) in (21), 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 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,

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

■

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

$$w(y(\iota, g), g) = \frac{1}{2}\left[s(\tilde{\iota}) + \frac{\alpha^{1-2b}\tilde{\iota}(g - \alpha\tilde{\iota})}{s(\tilde{\iota})}\right].$$

The truth-telling condition is

$$\iota \in \arg\max_{\hat{\iota}} \frac{1}{2}\left[s(\tilde{\iota}) + \frac{\alpha^{1-2b}\tilde{\iota}(g-\alpha\tilde{\iota})}{s(\tilde{\iota})}\right] - c(\iota, y(\hat{\iota}, g))$$

Notice that $\frac{\iota\eta}{\iota^b\eta^{1-b}} = f(\eta, \iota)$ .As in the proof of lemma 1, define $\hat{U}(\hat{\iota}, \iota)$ as

$$\hat{U}(\hat{\iota}, \iota) = \frac{1}{2}\left[f(\hat{\iota}, \hat{\eta}) + \alpha^{1-2b} f(\hat{\eta}, \hat{\iota})\right] - c(\iota, y(\hat{\iota}, g)),$$

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

Thus, the truth-telling condition can be stated as

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

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

$$\frac{1}{2}\left\{f_1(\iota, \eta) - \alpha f_2(\iota, \eta) + \alpha^{1-2b} f_2(\eta, \iota) - \alpha^{2-2b} f_1(\eta, \iota)\right\} - c_y(\iota, y(\iota, g))\, y_\iota(\iota, g) = 0. \tag{22}$$

The equation above can be simplified to:

$$y_\iota(\iota, g) = \frac{1}{2}\left\{(bg - \alpha\iota) f(\iota, \eta) + \alpha^{1-2b}\left[(1-b) g - \alpha\iota\right]\right\}.$$

Differentiating equation (22), we get

$$c_{y\iota}(\iota, y(\iota, g))\, y_\iota(\iota, g) = \frac{1}{2}\left\{\begin{array}{c} f_{11}(\iota, \eta) - 2\alpha f_{12}(\iota, \eta) + \alpha^2 f_{22}(\iota, \eta) + \\ \alpha^{3-2b} f_{11}(\eta, \iota) - 2\alpha^{2-2b} f_{12}(\eta, \iota) + \alpha^{1-2b} f_{22}(\eta, \iota) \end{array}\right\} \tag{23}$$

$$- c_{yy}(\iota, y(\iota, g))\left[y_\iota(\iota, g)\right]^2 - c_y(\iota, y(\iota, g))\, y_{\iota\iota}(\iota, g)$$

The second-order condition is

$$\begin{array}{c} \frac{1}{2}\left\{\begin{array}{c} f_{11}(\iota, \eta) - 2\alpha f_{12}(\iota, \eta) + \alpha^2 f_{22}(\iota, \eta) + \\ \alpha^{3-2b} f_{11}(\eta, \iota) - 2\alpha^{2-2b} f_{12}(\eta, \iota) + \alpha^{1-2b} f_{22}(\eta, \iota) \end{array}\right\} \\ - c_{yy}(\iota, y(\iota, g))\left[y_\iota(\iota, g)\right]^2 - c_y(\iota, y(\iota, g))\, y_{\iota\iota}(\iota, g) \leq 0. \end{array} \tag{24}$$

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

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

■

Proof of Proposition 1:

(i) Separating set: From equation (4), it follows that

$$w_y(y, g)\, y_\iota = s_\iota,$$

where we omit the dependence on (ι, g) for clarity.

Substituting y_ι from (7),

$$w_y = \frac{s_\iota}{s(bg - \alpha\iota)} = \frac{1}{\left[\iota\left(g - \alpha\iota\right)\right]^2}.$$

Continuous pooling set: From equation 10,

$$\frac{\partial w\left(\bar{y}\left(g\right), g\right)}{\partial g} = \frac{1-b}{\iota_b - \iota_a} M\left(\iota_b, \iota_a, g\right).$$

Discrete pooling set: From the firm's zero-profit condition, it follows that

$$w\left(y\left(\iota, g\right), g\right) = \frac{1}{2}\left[\iota^b\left(g - \alpha\iota\right)^{1-b} + \alpha^{1-2b}\iota^{1-b}\left(g - \alpha\iota\right)^b\right].$$

Differentiating the equation above, we get

$$\frac{\partial w}{\partial g} = \frac{1}{2}\left[\frac{(1-b)\,s\left(\iota\right) + b\alpha^{1-2b} f\left(\eta, \iota\right)}{\eta}\right] > 0.$$

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

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

But this is not truth-telling since

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

■

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) Clearly, a necessary condition for truth-telling is

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

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,

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

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

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

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

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

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:

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

(ii) A necessary condition for truth-telling is

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

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

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

Thus, the inequality can be written as

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

Analogously, the last individual in the discrete pooling must not want to get the contract of the first individuals in the separating set:

$$\hat{U}(\iota, \iota) \geq \lim_{x \to \iota_+} \hat{U}(x, \iota).$$

Hence, it follows that

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

implying in

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

which concludes the proof.

■

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 $[\iota_0, \gamma(\iota_0, g)]$ 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 13, 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 $\gamma(\iota_1, g) < \iota_0$, 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. ■

Proof of Lemma 8:

Assume that a type-(ι, g) strictly prefers to announce $\tilde{\iota} \neq \iota$:

$$\hat{U}(\tilde{\iota},\iota) \geq \hat{U}(\iota,\iota).$$

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

$$\begin{aligned}
\int_\iota^{\hat{\iota}} \left[\hat{U}_1(x,\iota) - \hat{U}_1(x,x)\right] dx &= \int_\iota^{\hat{\iota}} \int_x^{\iota} \hat{U}_{12}(x,z)\,dz dx \\
&= -\int_\iota^{\hat{\iota}} \int_x^{\iota} y_\iota(z,g)\, c_{y\iota}(z,g,y(x,g))\,dz dx > 0.
\end{aligned}$$

If $\hat{\iota}, \iota$ belong to a separating set, then this equation contradicts equation 7 (since x is between ι and $\hat{\iota}$).

Suppose $\hat{\imath}, \iota$ belong to a discrete pooling set (with no loss of generality, assume that $\hat{\imath} > \frac{g}{2\alpha} > \iota$). Then,

$$\hat{U}(\tilde{\imath}, \iota) = \hat{U}(\gamma(\tilde{\imath}, g), \iota), \text{ for all } \tilde{\imath}, \iota.$$

Thus, it follows that

$$\begin{aligned} \hat{U}(\hat{\imath}, \iota) - \hat{U}(\iota, \iota) &= \hat{U}(\hat{\imath}, \iota) - \hat{U}(\gamma(\iota, g), \iota) + \hat{U}(\gamma(\iota, g), \iota) - \hat{U}(\iota, \iota) \\ &= \hat{U}(\hat{\imath}, \iota) - \hat{U}(\gamma(\iota, g), \iota). \end{aligned}$$

Hence,

$$\hat{U}(\hat{\imath}, \iota) - \hat{U}(\iota, \iota) = -\int_{\gamma(\iota,g)}^{\hat{\imath}} \int_{x}^{\gamma(\iota,g)} y_\iota(z, g)\, c_{y\iota}(z, g, y(x, g))\, dz dx > 0.$$

But this contradicts equation 12 since x is between $\gamma(\iota, g)$ and ι.

Suppose $\hat{\imath}$ belongs to a separating set but ι belongs to a discrete pooling set. With no loss of generality, assume $\iota > \frac{g}{2\alpha}$. Then, it follows that

$$\begin{aligned} \hat{U}(\hat{\imath}, \iota) - \hat{U}(\iota, \iota) &= \hat{U}(\hat{\imath}, \iota) - \hat{U}(\gamma_0, \iota) + \hat{U}(\gamma_0, \iota) - \hat{U}(\iota, \iota) \\ &= \left[\hat{U}(\hat{\imath}, \iota) - \lim_{x \to \gamma_{0+}} \hat{U}(x, \iota)\right] + \left[\lim_{x \to \gamma_{0+}} \hat{U}(x, \iota) - \hat{U}(\gamma_0, \iota)\right] \\ &\quad + \left[\hat{U}(\gamma_0, \iota) - \hat{U}(\iota, \iota)\right]. \end{aligned}$$

The expressions in the first and last parenthesis are negative. Expanding the expression in the second parenthesis gives

$$\begin{aligned} \lim_{x \to \gamma_{0+}} \hat{U}(x, \iota) - \hat{U}(\gamma_0, \iota) &= \left[s(\gamma_0) - \frac{\lim_{x \to \gamma_{0+}} y(x)}{\iota(g - \alpha\iota)}\right] - \\ &\quad \left[\frac{s(\iota_0) + s(\gamma_0)}{2} - \frac{y(\gamma_0)}{\iota(g - \alpha\iota)}\right] \\ &= \frac{s(\gamma_0) - s(\iota_0)}{2} + \frac{y(\gamma_0) - \lim_{x \to \gamma_{0+}} y(x)}{\iota(g - \alpha\iota)} \end{aligned}$$

Substituting the equation from Condition 4 implies in

$$\begin{aligned} \lim_{x \to \gamma_{0+}} \hat{U}(x, \iota) - \hat{U}(\gamma_0, \iota) &= \frac{s(\gamma_0) - s(\iota_0)}{2} + \frac{\iota_0(g - \alpha\iota_0)}{\iota(g - \alpha\iota)} \frac{s(\iota_0) - s(\gamma_0)}{2} \\ &= \frac{s(\gamma_0) - s(\iota_0)}{2}\left[1 - \frac{\iota_0(g - \alpha\iota_0)}{\iota(g - \alpha\iota)}\right] \\ &= \frac{s(\gamma_0) - s(\iota_0)}{2}\left[1 - \frac{\gamma_0(g - \alpha\gamma_0)}{\iota(g - \alpha\iota)}\right]. \end{aligned}$$

As $\gamma_0 > \iota > \frac{g}{2\alpha}$, it follows that $\frac{\gamma_0(g-\alpha\gamma_0)}{\iota(g-\alpha\iota)} < 1$. Hence, $\lim_{x \to \gamma_{0+}} \hat{U}(x, \iota) - \hat{U}(\gamma_0, \iota) < 0$. Thus, $\hat{U}(\hat{\imath}, \iota) - \hat{U}(\iota, \iota) < 0$ (where $s(\gamma_0) - s(\iota_0) < 0$ since the technology is intensive in cognitive skills). ■

Proof of Lemma 9:

Define $\hat{U}(\hat{\iota},\iota)$ as the utility received by a type-(ι,g) individual who gets a contract designed for a type $(\hat{\iota},g)$ individual. Then, from the truth-telling condition, we get the following local first- and second-order conditions:

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

The first-order condition yields, for all ι,

$$s_\iota(\iota,y)+[s_y(\iota,y)-c_y(\iota,g,y(\iota,g))]\,y_\iota(\iota,g)=0\therefore \tag{26}$$

$$y_\iota(\iota,g)=\frac{s_\iota(\iota,y)}{c_y(\iota,g,y(\tilde{\iota},g))-s_y(\iota,y)}. \tag{27}$$

Simplifying the equation above, we get

$$y_\iota(\iota,g)=s(y,\tilde{\iota})\,y(\iota,g)\,\frac{bg-(1-a)\,a\iota}{y(\iota,g)-as(y,\tilde{\iota})}.$$

Analogously to the proof of Lemma 1, differentiating equation (26) and substituting in the second-order condition, yields $y_\iota(\iota,g)(g-2\alpha\iota)\geq 0$. ∎

Proof of Lemma 10:

Define $\hat{U}(\hat{\iota},\iota)$ as the utility received by a type-(ι,g) individual who gets a contract designed for a type $(\hat{\iota},g)$ individual and notice that $y^a\eta^b\iota^{1-a-b}=f(\eta,\iota,y)$. Then, from the truth-telling condition, the local first-order condition is

$$y_\iota=y\,\frac{f(\iota,\eta,y)\,[b\eta-(1-a-b)\,\alpha\iota]+f(\eta,\iota,y)\,\alpha^{1-a-2b}\,[(1-a-b)\,\eta-bg]}{y(\iota,g)-a\iota\eta\,[f(\iota,\eta,y)+\alpha^{1-a-2b}f(\eta,\iota,y)]}.$$

Analogously to the proof of Lemma 1, differentiating the equation above and substituting in the second-order condition, yields $y_\iota(\iota,g)(g-2\alpha\iota)\geq 0$. ∎

References

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

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

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

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

[5] 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

[19] 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.

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

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

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

[23] 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.

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

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

[26] 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.

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

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

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

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

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

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