RAND Journal of Economics
Vol. 38, No. 4, Winter 2007
pp. 1020–1043

A model of mixed signals with applications to countersignalling

Aloisio Araujo*

Daniel Gottlieb**

and

Humberto Moreira***

We develop a job-market signalling model where signals convey two pieces of information. This model is employed to study countersignalling (signals nonmonotonic in ability) and the GED exam. A result of the model is that countersignalling is more likely 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 noncognitive skills and (ii) it does not affect wages.

1. Introduction

■ In the initial papers in the signalling literature, the informational asymmetry consists of a unidimensional parameter which is known to only one side of the market (e.g., Spence, 1973). Then, under the natural condition that individuals can be ordered according to their marginal utility of signalling (single-crossing property), there exists a family of separating equilibria in which signals reveal information monotonically. In job-market models, for example, higher education discloses information about higher productivity. These equilibria are ranked by the Pareto optimality criterion; moreover, only the Pareto dominant equilibrium is robust to competition among firms (Riley, 1979).

More recently, scholars have identified conditions under which the main results from unidimensional models extend to multidimensional ones. These conditions typically involve

* IMPA and Getulio Vargas Foundation; aloisio.araujo@fgv.br.
** Massachusetts Institute of Technology; daniel_g@mit.edu.
*** Getulio Vargas Foundation; humberto@fgv.br.
Part of this research was done when the second author was a student at the Getulio Vargas Foundation. The authors thank Juliano Assunção, Abhijit Banerjee, Luis Henrique Braido, Flávio Cunha, Carlos E. da Costa, Mathias Dewatripont, Glenn Ellison, James Heckman, Bengt Holmstrom, Derek Neal, Jean Tirole, Marcos Tsuchida, a coeditor, and two referees for helpful comments and suggestions. Earlier versions of this article were presented in the Meeting of the Brazilian Econometric Society (João Pessoa, 2004), the Conference in Tribute to Jean-Jacques Laffont (Université de Toulouse, 2005), the Summer Meeting of the Econometric Society (Minneapolis, 2006), the Latin American Meeting of the Econometric Society (Mexico, 2006), the Workshop in Mathematical Economics (Rio, 2006), and seminars at the University of Chicago, the Getulio Vargas Foundation, and MIT. All remaining errors are the authors' alone.

some form of separability between dimensions. However, in many situations, a single instrument conveys information about multiple characteristics. In such cases, good and bad characteristics may be revealed by the same instrument. We refer to a multidimensional model with a single signalling instrument as a "mixed-signals" model.[1] Another strand of literature has focused on unidimensional models featuring nonmonotonic signals. These models generate "countersignalling," wherein individuals with high types choose to engage in a lower amount of signalling than medium-type individuals (Feltovich, Harbaugh, and To, 2002).

In this article, we present a two-dimensional characteristics signalling model satisfying the single-crossing property (SCP) in each dimension. Workers' characteristics are represented by a vector of cognitive and noncognitive ability parameters. Firms can observe a combination of these characteristics through an interview but cannot precisely determine whether the result of this interview was because of high cognitive or noncognitive ability. The results of the interview process can be used to reduce the two-dimensional model to a one-dimensional model where the SCP fails to hold. Workers are able to signal their characteristics through the number of years of education they acquire. A theoretical contribution of the article is to provide a characterization of the equilibrium in a signalling model where the SCP fails, thereby extending Araujo and Moreira's (2001) analysis of a screening model.

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

This model is also employed in order to understand evidence on the General Educational Development (GED) exam. The signalling equilibrium has some interesting properties consistent with available empirical evidence on the GED: individuals with different abilities obtain the same amount of education, and passing the exam signals higher cognitive skills but does not increase one's earnings. These results follow from the fact that the GED is a mixed signal: if a worker with low overall ability has passed the exam, it means that her noncognitive 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 ineffectiveness of the GED exam stems from its focus on cognitive ability. A test that places a stronger emphasis on noncognitive 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.

□ **Related literature on countersignalling and mixed signals.** There are several documented examples of what appears to be countersignalling behavior. Hvide (2003) argues that intermediate individuals appear to pursue more education than bright individuals for professions where individuals without a license are not denied work. Unlike standard signalling models of advertising predict, Clements (2004) documents that many high-quality products are sold in low-quality packages. Orzach, Overgaard, and Tauman (2002) suggest that, even controlling for market size, luxury cars (such as Rolls Royce and Ferrari) seem to be advertised very modestly compared to nonluxury cars.[2] In the context of fashion as a signal of status, Pinker (1999) claims that "trend-setters are members of upper classes who adopt the styles of lower classes to differentiate themselves from middle classes." According to O'Neil (2002), countersignalling led intermediately advanced countries to spend more on their military than most advanced countries after World War II.

Feltovich, Harbaugh, and To (2002) present a countersignalling model applied to the labor market. Firms access some measure of the worker's ability (which is interpreted as the

[1] We have borrowed the term "mixed signal" from Cavallo, Heckman, and Hsee (1998).

[2] Caves and Greene (1996) find no significant systematic positive correlation between quality and the amount of advertising.

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 a schooling activity. If the exogenous signal were sufficiently informative about the individual productivity of workers, then it would not be profitable for them to use schooling to signal productivity. On the other hand, if the exogenous signal were completely uninformative about the workers' productivities, we would have a standard signalling model where higher types signal more. Their model can be seen as an intermediate case where the exogenous signal is sufficiently informative to separate high from medium types but not sufficient to separate medium from low types.

Our model differs from that of Feltovich, Harbaugh, and To in that uncertainty about productivity comes from the divergence between the schooling technology and the firms' technology instead of a noise term. The misalignment between these two technologies generates an incentive for some higher-productivity workers to obtain less education.

Orzach, Overgaard, and Tauman (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 (i.e., cuts below the full-information level of the low-quality firm) 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.

One example of mixed signals is the GED exam, which is taken by high school dropouts to certify that they have knowledge equivalent to high school graduates. The GED reveals, at the same time, high cognitive skills and low noncognitive skills (Cameron and Heckman, 1993; Cavallo, Heckman and Hsee, 1998; Heckman and Rubinstein, 2001). Moreover, wages received by high school dropouts are not influenced by this certificate. Another example is the so-called "PhD curse."[3] This curse refers to the difficulty of some recent PhD graduates to find jobs outside of academia because firms perceive them as being too theoretically oriented and lacking enough practical abilities.[4] A third example of mixed signals is presented by Drazen and Hubrich (2003), who argue that higher interest rates show that the government is committed to maintaining a fixed exchange rate, but also signals weak fundamentals. Benabou and Tirole (2006) argue that donations signal altruism, but may also signal a desire to be perceived as altruistic.

In the labor-market model, for example, an assumption of unidimensional information asymmetry implies that all relevant characteristics of an employee can be captured by a single ability-type, usually thought of as cognitive ability. However, significant empirical evidence supports the importance of noncognitive skills as well as cognitive skills in the labor market (Heckman, Stixrud, and Urzua, 2006). Generalization of the original results to the multidimensional case turns out not to be straightforward; in Kholleppel's (1983) example of a two-dimensional extension of Spence's model, no separating equilibrium exists.

Quinzii and Rochet (1985) and Engers (1987) provide sufficient conditions for the existence of a separating equilibrium in the multidimensional model. In Quinzii and Rochet, ability was represented by a k-dimensional vector and they assume the existence of k (nonexclusive) different types of education. Moreover, they assume that the signalling costs were linear and separable in the signals (up to a change of variables). Hence, it was as though 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 (noncognitive skills). Under this separability assumption (which implies that the single-crossing property holds in each dimension), Quinzii and Rochet obtain results similar to the unidimensional model: separating equilibria exist and wages are monotonic in the worker's types.

Needless to say, the educational systems assumed by Quinzii and Rochet are not realistic because all known educational systems require both cognitive and noncognitive abilities (although

[3] We thank Mathias Dewatripont for this example.

[4] See "Academic Careers: A Comparative Perspective," Jeroen Huisman and Jeroen Bartelse, (eds.), available from www.awt.nl/uploads/files/academic.pdf.

in different proportions). Engers relaxes this assumption through a generalization of the unidimensional assumption that individuals' marginal utility of signalling could be ordered (single-crossing property). However, in the multidimensional case, this assumption is much less compelling because, as the number of signals rises, it becomes more likely that the SCP will not hold when one controls for one signal (i.e., the introduction of other signals may break the SCP in the multidimensional case).

The rest of the article is organized as follows. The basic framework is presented in Section 2. Section 3 characterizes the equilibrium. Section 4 studies how countersignalling may emerge and Section 5 employs this framework to analyze the GED exam. Section 6 briefly discusses examples of other environments where the model can be applied. Then, Section 7 concludes.

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 and η represent cognitive and noncognitive ability, respectively. For concreteness, we will refer to ι as intelligence and η as perseverance while bearing in mind that noncognitive skills embody several other characteristics such as motivation, self-control, and other personality traits. The set of all possible characteristics is the compact set $\Theta \equiv [\iota_0, \iota_1] \times [\eta_0, \eta_1] \subset \mathbb{R}^2_{++}$ and the types are distributed according to a continuous density $p : \Theta \to \mathbb{R}_{++}$, which is assumed to be a C^2 function.

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

Firms have identical technologies with constant returns to scale $f(\iota, \eta)$ and act competitively.[5] Moreover, other than schooling, firms observe the result of an interview $g(\iota, \eta)$ which does not fully reveal the worker's productivity. Thus, even though firms have some idea of the overall ability of a worker, they are unable to unambiguously determine her productivity by observing the result of the interview.[6] In a more general model, we could imagine that individuals might exert effort in order to distort the market's assessment of their productivity (e.g., Holmstrom, 1999; Dewatripont, Jewitt, and Tirole, 1999). We studied this possibility in a previous version of the article, where we assumed that schooling influences the worker's performance in the interview. Most of the results in this article are unaffected.[7]

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 signalling 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).

Because firms are homogeneous, 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 signalling game is a profile of strategies $\{y(\iota, g), w(y, g)\}$ and beliefs $\mu(\cdot \mid y, g)$ such that

[5] In this article, we consider only the pure signalling case. In a previous version of the article, we have shown that all the results also hold when schooling affects productivity (see Araujo, Gottlieb, and Moreira, 2007).

[6] The hypothesis that firms can access an additional signal that consists of a measure of the worker's ability is also present at Feltovich, Harbaugh, and To (2002).

[7] See Araujo, Gottlieb, and Moreira (2007). In this case, it can be shown that, locally, the ability to distort the result of the interview raises the amount of education in equilibrium for all individuals as in standard "signal-jamming" models.

(i) The worker's strategy is optimal given the equilibrium wage schedule:

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

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

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

Next, we will specify the analytical forms of the functions presented.[8] The signalling technology is characterized by the following cost of signalling function:

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

The function above implies that the cost of education is decreasing in intelligence and perseverance. Moreover, intelligence and perseverance are imperfect substitutes in the schooling process.[9]

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. Notice that, conditional on the interview g, the workers' types lie on the hyperplane represented by equation (2). Hence, by conditioning on g, the type space becomes one-dimensional. Thus, we will refer to a type-(ι, η) worker as type-$(\iota;g)$, as ι captures all private information after taking g into account.

Substituting (2) into (1), we are able to rewrite the cost of signalling 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.

In general, the single-crossing property (SCP) may not be satisfied because

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

The SCP states that the marginal utility of effort is monotonic in the worker's type. 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. In particular, even for individuals with very low intelligence and very high perseverance levels, raising a unit of intelligence and decreasing α units of perseverance would decrease the marginal cost of schooling (or the opposite when the sign of $c_{y\iota}$ is reversed). 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 into 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 signalling given the interview result g. When the SCP is satisfied, $[\iota_0, \iota_1]$ is contained in one of these intervals. Figure 1 depicts a situation where $[\iota_0, \iota_1]$ overlaps these intervals.

Therefore, the two-dimensional model reduces to a one-dimensional model where the SCP may fail to hold after conditioning on g.

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

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

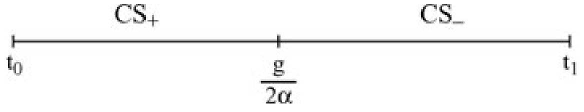
where $b \in (0, 1)$. If $b > \frac{1}{2}$ we say that the firm's technology is intensive in cognitive skills. Otherwise, we say that it is intensive in noncognitive skills.

[8] The robustness of the model to the functional forms is studied in Appendix A.

[9] They are "cost substitutes" in the sense that $c(\iota, \eta, y)$ has increasing differences.

FIGURE 1

CS+ AND CS− INTERVALS



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

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

3. The signalling equilibria

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

Given an equilibrium profile of education $y(\iota, g)$, we refer to the set of types whose signal is (y, g) as the pooling set $\Theta(y, g)$. A type $(\iota;g)$ is separated if, in equilibrium, her characteristics are revealed from her signals $y(\iota, g)$ and g. If more than one type chooses the same amount of education, we say that they are pooled.

In signalling models where the SCP is satisfied, incentive compatibility requires education to be increasing if $c_{y\iota} < 0$ for all ι (CS_+) and decreasing if $c_{y\iota} > 0$ for all ι (CS_-). Then, if two workers are pooled, monotonicity implies that all intermediate types must also pool with them (see graph on the left in Figure 2). We call these types *continuously pooled*.

When the single-crossing property does not hold, incentive compatibility requires y to be increasing in the CS_+ interval and decreasing in the CS_- interval so that the equilibrium may feature nonmonotone signalling. As a result, a disconnected set of workers may acquire the same level of education (see graph on the right in Figure 2). We say that these workers are *discretely pooled*.[10]

The precise definitions are stated below:

Definition 2. Given an equilibrium profile of education $\{y(\iota, g) : \iota \in [\iota_0, \iota_1], g \in [\alpha\iota + \eta_0, \alpha\iota + \eta_1]\}$:

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

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

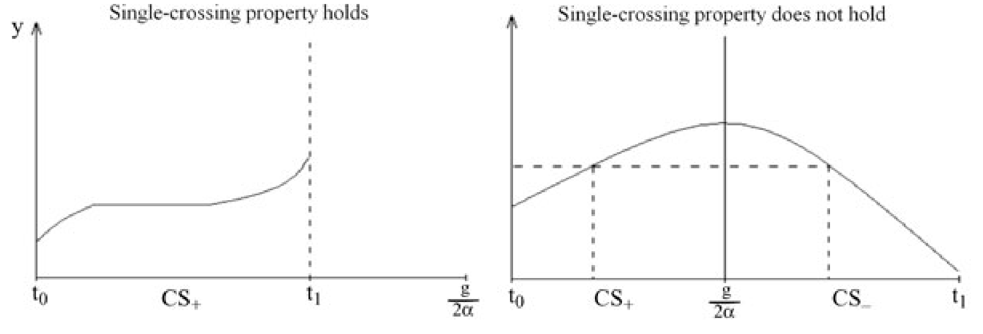
(iii) A type$-(\iota;g)$ worker is discretely pooled if $\Theta(y(\iota, g), g) \neq \{(\iota, g)\}$ is discrete. A discrete pooling set is a set of types where every element is discretely pooled.

In any signalling equilibrium, each type must belong to one of these three sets. The selection criterion (to be discussed later) will rule out continuous pooling. Hence, the equilibrium will take a form similar to the profile on the right of Figure 2: there will be an interval of separated types and an interval of discretely pooled types.

[10] Analyzing a model of competition in many markets, Green and Laffont (1990) have also obtained an equilibrium where discrete pooling may occur. When the incumbent is able to commit to her decisions, the incentive-compatibility constraints are relaxed so that the equilibrium may feature discrete pooling. In contrast, in our article, discrete pooling occurs under the standard incentive-compatibility constraints.

FIGURE 2

POOLING AND INCENTIVE-COMPATIBILITY



In the following subsections, we study the properties of separating sets, continuous pooling sets, and discrete pooling sets, respectively. As is standard in the signalling literature, the equilibrium will be represented by differential equations that follow from the worker's first-order condition. Hence, we characterize piecewise C^2 equilibria.[11]

□ **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

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

The worker's incentive-compatibility constraint is

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

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 the worker's productivity is maximal given the interview result g. This educational level is defined as

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

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 because a worker whose intelligence is too high must have a low level of perseverance.

As a worker must be earning her expected productivity in any separating set, however, it follows that wages are nonmonotone in intelligence (controlling for the interview g). As shown in the previous signalling literature, when the SCP is satisfied, education is increasing in the worker's type. Suppose this is also the case when the SCP is not valid (i.e., suppose that

[11] A piecewise C^k function is a function whose domain can be partitioned into a finite number of intervals such that the function is k times continuously differentiable in each interval. Therefore, we allow for a finite number of jumps and kinks. Our results can be generalized to piecewise C^1 functions. However, as most of the literature, we focus on the piecewise C^2 case in order to simplify the proofs.

education is increasing in intelligence). Then, firms would offer a higher salary for individuals with intermediate schooling (as those are the most productive workers).[12] Such an allocation cannot, however, be an equilibrium because the 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 incentive-compatibility is that education must be increasing in ι until $\iota^*(g)$ and decreasing after $\iota^*(g)$.

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

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

which implies that y must be increasing if $\iota < \iota^*(g)$ and 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 program in equation (5) we obtain the following lemma, whose proof is presented in the Appendix.

Lemma 1. In any separating set, if an education and wage profile is incentive compatible, it must satisfy

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

and equation (7). Moreover, in a separating set, the workers with the highest level of schooling are the most productive (not the brightest or the most perseverant) and the level of schooling is strictly increasing in productivity.

Remark 1. Notice that equation (7) implies that

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

Moreover, equation (8) yields

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

Generally, equations (10) and (9) cannot hold for all ι except when $b = \frac{1}{2}$. In this case, the firms' technology is identical to the signalling technology. If so, 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 because the relative intensity of intelligence in the schooling technology is different from that in the firm's technology. Then, if either $\frac{bg}{\alpha} \in [\iota_0, \iota_1]$ or $\frac{g}{2\alpha} \in [\iota_0, \iota_1]$,there must exist some pooling in equilibrium (as the separating set conditions cannot hold for all the intervals of parameters).

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

The zero-profit condition is

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

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

where $W(I, g) \equiv \int_I f(x, g - \alpha x)p(x \mid g)\,dx$ is the expected productivity of a type-ι worker. Conditions ii and iii from Definition 1 are trivially satisfied in that given set.

□ **Discrete pooling set.** A distinct feature of models where the SCP does not hold is the emergence of discrete pooling, where individuals with nonadjacent types receive the same contract. This feature is a direct consequence of the possibility of nonmonotone signals.

As was shown by Araujo and Moreira (2001), a necessary condition for incentive compatibility in a discrete pooling set is the so-called marginal utility identity, according to which, if two individuals are (discretely) pooling in a contract, they should have the same marginal utility. We formally state that result as a lemma:

Lemma 2. Suppose that $\{y(\iota, g) : \iota \in [\iota_0, \iota_1], g \in [\alpha\iota + \eta_0, \alpha\iota + \eta_1]\}$ is an incentive-compatible profile of education. If two regular workers with the same interview result choose the same level of education, then their marginal cost of education must be the same:

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

The economic interpretation of Lemma 2 is that if two nonadjacent 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 signalling, it follows that if a type-(ι;g) worker is in a discrete pooling set, the other worker pooling with her is ($\hat{\iota}$; g) defined as

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

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

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

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

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

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

Analogously to Lemma 1, the local first- and second-order conditions from the workers' incentive-compatibility constraint yield the following.

Lemma 4. If (ι;g) belongs to a discrete pooling set, then if an education and wage profile is incentive compatible, they satisfy

$$\begin{aligned} y_\iota(\iota, g) = {} & s(\iota, g)[P(\iota, g)(bg - \alpha\iota) + P_\iota(\iota, g)\,\iota(g - \alpha\iota)] \\ & + s\left(\frac{g}{\alpha} - \iota, g\right)\{[1 - P(\iota, g)][(1 - b)g - \alpha\iota] + P_\iota(\iota, g)\,\iota(g - \alpha\iota)\}, \end{aligned} \tag{13}$$

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

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

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

□ **Equilibrium selection and comparative statics.** The proposition below presents some comparative statics results. Because education is costly, a worker would only choose to obtain an additional amount of education if it increases wages. Thus, incentive compatibility requires wages to be strictly monotonic.

Proposition 1. In any PBE, wages are strictly increasing and concave in the amount of schooling controlling for the interview.

Notice that, for fixed ι, the productivity is increasing in the result of the interview g. Then, in a separating set, wages must be increasing in g. However, this may not be true in a pooling set: because wages are a weighted average of the productivity of pooled types (where weights are given by the conditional probability of each type), a change in g would also affect the weights attributed to each type. In a discrete pooling set, for example, it follows that[14]

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

The first and second terms are positive and represent the direct effect: more productive individuals obtain a higher result in the interview. The last term may be either positive or negative and reflects the indirect effect. If the proportion of more productive individuals is decreasing in g, then this term is negative.[15] If $\iota \mid g$ is uniformly distributed, for example, then this last term vanishes (because the conditional distribution of ι is constant), implying that wages are increasing in the interview.

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

As the concept of PBE leads to multiple equilibria, we will apply a selection criterion in order to pick an equilibrium. Riley (1979) suggests 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 signalling literature.

As a fully separating equilibrium does not exist when the single-crossing property does not hold, one must employ a weaker refinement criterion. We propose 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-separability criterion selects the most efficient equilibrium in the class of equilibria with the highest degree of separation.

Definition 3. A quasi-separable equilibrium is a PBE that satisfies the following conditions:

(i) If two workers belong to a pooling set, then their marginal cost of schooling must be the same.

(ii) There is no other PBE satisfying Condition i such that every type obtains less schooling (with strictly less for at least one type).

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

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

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

The first condition identifies the highest possible degree of separability by ruling out continuous pooling. The second condition gives the boundary condition that uniquely determines the equilibrium. It consists of a Pareto improvement criterion for selection.

The following proposition can be seen as evidence that the SCP does not hold. It states that two individuals with different abilities obtaining the same amount of schooling are 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 there exists pooling in a quasi-separable equilibrium, then the SCP does not hold.

□ **Characterization of the equilibrium.** In this subsection, we characterize the equilibrium of the model. As the results are more technical than the rest of the article, readers more interested in the applications of the model can skip this subsection.

As in equation (12), we denote by $\gamma(\iota, g)$ the type with the same marginal cost of signalling as $(\iota;g)$. 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). Clearly, as $\gamma(\iota_0, g) \leq \iota_1$, 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 $\iota \in (\gamma(\iota_0, g), \iota_1])$. In this subsection, we show that the quasi-separable equilibrium is such that all types outside of this interval are discretely pooled (where a pool consists of two nonadjacent types). The characterization is carried out through a series of lemmata.

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

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

The first lemma establishes another necessary condition for incentive compatibility.

Lemma 5. $U(\iota, \cdot, g)$ is continuous for all $\iota \in [\iota_0, \iota_1]$.

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

The continuity of U enables us to determine the boundary condition for the amount of education when switching from discrete pooling sets to separating sets. Notice that when a worker becomes pooled with another type, his expected productivity changes discontinuously (as it becomes the average of their productivities). Thus, the wage function has a discontinuity when switching from separating sets to discrete pooling sets. Hence, the education must be discontinuous in order to preserve the continuity of the indirect utility. This is formally established in the following corollary.[16]

Corollary 1. Let ι be such that $[\iota - \varepsilon, \iota]$ is a discrete pooling set and $(\iota, \iota + \varepsilon]$ is a separating set, for some $\varepsilon > 0$. If $y(\cdot, g)$ is right continuous at ι, the following condition is necessary for incentive compatibility:

$$y(\iota, g) = \iota(g - \alpha\iota)[1 - P(\iota, g)]\,[s(\gamma(\iota, g), g) - s(\iota, g)] + \lim_{x \to \iota_+} y(\iota, g). \tag{15}$$

From now on we assume that the education profile is right continuous. The second lemma determines the maximal discrete pooling set.

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

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 into two intervals: a discrete pooling interval $[\iota_0, \gamma(\iota_0, g)]$ and a separated interval $(\gamma(\iota_0, g), \iota_1]$.

[16] We obtain the same result if $[\iota, \iota + \varepsilon]$ is a discrete pooling set and $[\iota - \varepsilon, \iota]$ is a separating set for some $\varepsilon > 0$.

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 $(\iota_1 ;g)$ 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). However, this would also reduce the cost of schooling. Thus, in equilibrium, ι_1 must choose the lowest amount of schooling possible.

The next proposition establishes that the conditions from Lemmata 1, 4, and 7 and Corollary 1 are also sufficient for the quasi-separable equilibrium.

Proposition 3. A profile of education is a quasi-separable equilibrium if, and only if, it satisfies:

(1) $y_\iota(\iota, g) = s(\iota, g) \times (bg - \alpha\iota)$ for $\iota > \gamma(\iota_0)$,
(2) $y(\iota_1, g) = 0$,
(3) $y_\iota(\iota, g) = s(\iota, g)\,[P(\iota, g)(bg - \alpha\iota) + P_\iota(\iota, g)\iota(g - \alpha\iota)] + s(\frac{g}{\alpha} - \iota, g)\{[1 - P(\iota, g)]\,[(1-b)g - \alpha\iota] + P$ for $\iota < \gamma(\iota_0)$,
(4) $y(\gamma(\iota_0, g), g) = P(\iota_0, g)[s(\gamma(\iota_0, g), g) - s(\iota_0, g)]\iota_0(g - \alpha\iota_0) + \lim_{x\to\gamma(\iota_0)-} y(x, g)$.

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.

The amount of education for separated types is determined from condition (1) of Proposition 3 and boundary condition (2). Then, using conditions (3) and (4) (a differential equation with a boundary condition), one can calculate the equilibrium amount of education for discrete pooling types.

Notice that item 4 from Proposition 3 implies that education must jump downward at $\gamma(\iota_0, g)$ as $s(\iota_0, g) - s(\gamma(\iota_0, g), g) > 0$ (see Lemma 3 and $b < 1/2$). This follows from the fact that wages are discontinuous: individuals with $\iota \in [\frac{g}{2\alpha}, \gamma(\iota_0, g)]$ earn wages higher than their productivity because they are pooled with more productive workers, but those with types higher than $\gamma(\iota_0, g)$ earn their productivity because they are separated. Hence, if education were continuous, indirect utility would be discontinuous. As shown in Lemma 5, however, a discontinuous indirect utility is not incentive compatible. Thus, the amount of education must jump downward in order to preserve the continuity of the indirect utility function.

The main form of the equilibrium can be captured by the following symmetric example. Suppose $\alpha = 1$ and $\Theta = [1, 10]^2$. Then, $g \in [2, 20]$ and

$$c_{y\iota}(\iota, g, y)\begin{Bmatrix} > \\ < \end{Bmatrix} 0 \Leftrightarrow \iota \begin{Bmatrix} > \\ < \end{Bmatrix} \frac{g}{2}.$$

Consider different values of the interview result g. For the lowest possible value, $g = 2$, the worker's type is uniquely revealed ($\iota = \eta = 1$). Hence, there are no incentives for signalling so that $y(\iota, g) = 0$ for $\iota = 1, g = 2$. Analogously, if $g = 20$, the worker's type is uniquely revealed to be $\iota = \eta = 10$ and there are no incentives for signalling.

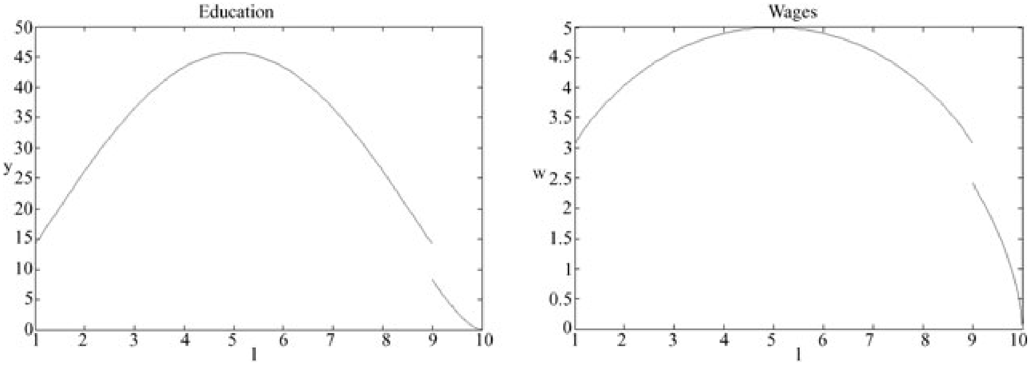
For $g \in (2, 20)$, the CS_+ and CS_- intervals are $[1, \frac{g}{2}]$ and $\left[\frac{g}{2}, 10\right]$, respectively. Education is increasing in ι for $\iota \leq \frac{g}{2}$ and decreasing for $\iota \geq \frac{g}{2}$. Because of the symmetry of the example, pairs equidistant from $\frac{g}{2}$ are discretely pooled and extreme types are separated. When $g \in (2, 11)$, the discrete pooling interval is $[1, g - 1]$, the separating interval is $[g - 1, 10]$, and y is discontinuous at $\iota = g - 1$. When $g \in (11, 20)$, the discrete pooling interval is $[g - 10, 10]$, the separating interval is $[1, g - 10]$, and y is discontinuous at $\iota = g - 10$.

Figure 3 presents the equilibrium amount of education and wages conditional on $g = 10$ for the case where $b = 0.4$, $\alpha = 1$, $\iota_0 = 1$, $\iota_1 = 10$, and $\iota \mid g \sim U[\iota_0, \iota_1]$.[17]

[17] In a previous version of the article, we have presented the equilibrium profiles of education, wages, and utility for other parameters. See Araujo, Gottlieb, and Moreira (2007).

FIGURE 3

EQUILIBRIUM IN THE SYMMETRIC MODEL (a)



In Figure 4, the equilibrium amount of utility and the profile of wages as a function of education conditional on $g = 10$ are presented. Notice that both education and wages are discontinuous but the utility is continuous in ι. As Proposition 1 shows, wages are strictly increasing and concave in education.

4. Countersignalling

■ In this section, we show how the basic model allows us to understand the existence of countersignalling. First, we present a precise definition.

Definition 4. A type-(ι;g) worker is countersignalling if

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

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

As shown in the last subsection of Section 3, education is strictly increasing for $\iota < \frac{g}{2\alpha}$ and strictly decreasing for $\iota > \frac{g}{2\alpha}$. Moreover, as argued in the first subsection of Section 3, the productivity of a worker with interview result g is strictly increasing for $\iota < \frac{bg}{\alpha}$ and strictly decreasing for $\iota > \frac{bg}{\alpha}$. Thus, the countersignalling interval is $[\frac{bg}{\alpha}, \frac{g}{2\alpha}]$. Hence, countersignalling occurs if, and only if, the schooling technology is different from the firms' technology, that is, $b \neq \frac{1}{2}$.

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 $|b - \tilde{b}|$. Then, the distance from the schooling technology to the firms' technology is given by $\frac{1}{2} - b$. Notice that increasing the distance between the two technologies (i.e., reducing b) strictly increases the countersignalling interval. Thus, we have proved the following.

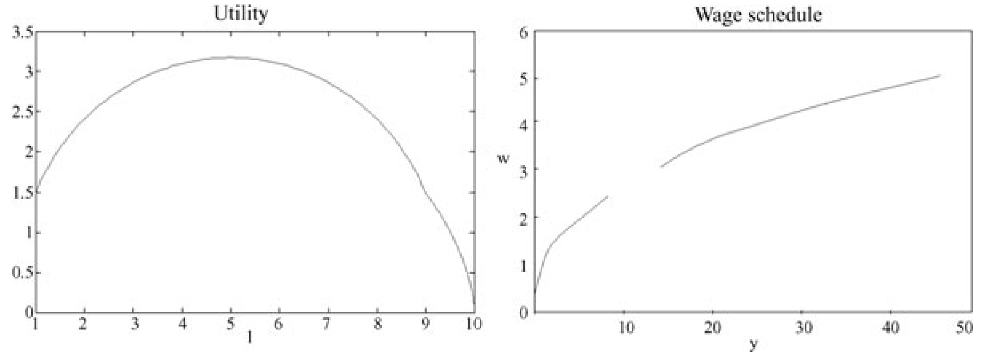
Proposition 4. Countersignalling occurs if and only if the schooling and the firms' technologies are different (i.e., the SCP does not hold), and the countersignalling interval strictly increases in the distance from the schooling technology to the firms' technology.

This proposition provides an intuitive testable implication. Countersignalling 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 levels of education should be more common among sportsmen and artists than among teachers.

The analysis above was made conditional on the interview g. It turns out that education and wages may also be nonmonotone in the abilities unconditionally. More specifically, consider the original (two-dimensional) type space. Define the education obtained by type-(ι, η) as

FIGURE 4

EQUILIBRIUM IN THE SYMMETRIC MODEL (b)



$$\tilde{y}(\iota, \eta) \equiv y(\iota, g(\iota, \eta)).$$

Suppose that $\Theta = [\theta_0, \theta_1]^2$. Then, the worker's type is uniquely revealed when $\iota = \eta = \theta_0$ so that she has no incentive to signal. Analogously, her type is uniquely revealed when $\iota = \eta = \theta_1$. Hence, $\tilde{y}(\theta_0, \theta_0) = \tilde{y}(\theta_1, \theta_1) = 0$ so that $\tilde{y}(\iota, \eta)$ cannot be monotone (as long as $\tilde{y}(\iota, \eta) > 0$ for some ι and η).[18]

The reason for this unconditional nonmonotonicity arises from the worker's incentive to signal. When ι and η are extreme, there is not much uncertainty regarding the individual's type. Thus, she faces low incentives to signal. However, when the worker has moderate types, there are many different types with the same interview g. Hence, she has high incentives to signal.[19]

A key message of the model is that, when types are multidimensional, signals that reveal a high type in one dimension may indicate a low type in other dimensions. This result captures the idea that an employer may be suspicious that a potential employee who looks "too perfect" on one dimension may have problems in other unobserved dimensions. One intuitive case occurs when the (two-dimensional) types are perfectly negatively correlated. Then, one could reparameterize the model into a one-dimensional type model where the SCP does not hold. In the model above, types are perfectly negatively correlated conditional on the interview g. This leads to the SCP not being satisfied and to the emergence of countersignalling.

5. The GED exam

■ **Empirical evidence.** Signalling models (e.g., Spence, 1973) generally assume that an individual's personal abilities are represented by a scalar of cognitive skills. However, a vast body of empirical evidence consistently contradicts this assumption. Heckman, Stixrud, and Urzua (2005), for example, find that for several dimensions of behavior and for a variety of labor-market outcomes, noncognitive skills are better predictors of behavior than cognitive skills.[20]

In the psychology field, the five-factor model of personality (referred to as the "Big Five") identifies five dimensions of noncognitive characteristics: extroversion, conscientiousness,

[18] When education is productive, $\tilde{y}(\theta_0, \theta_0)$ and $\tilde{y}(\theta_1, \theta_1)$ will generally not be zero.

[19] Therefore, as in Feltovich, Harbaugh, and To (2002), the presence of an additional signal makes intermediate types the ones with the highest incentives to signal.

[20] Cawley et al. (1996) showed that cognitive ability is only a minor predictor of social performance and that many noncognitive factors are important determinants of blue collar wages. Bowles and Gintis (2001) provide survey evidence that employers consider measures of noncognitive skills to be significantly more important than measures of cognitive skills in the hiring of production workers. Klein, Spady, and Weiss (1991) show that lower quit rates and lower absenteeism account for most of the premium awarded by high school graduates compared to high school dropouts (*not* higher productivity). Edwards (1976) shows that dependability and consistency are more valued by blue collar supervisors than cognitive ability and independent thought.

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).

An interesting set of evidence on the impact of noncognitive skills on education and wages comes from the General Educational Development (GED). The GED is an exam taken by American high school dropouts to certify that they have equivalent knowledge to high school graduates. It started in 1942 as a way to allow veterans without a high school diploma to obtain a secondary school credential. Today, about half of the students who drop out of high school and a fifth of those classified as "high school graduates" by the U.S. Census Bureau are GED recipients.

The GED consists of five tests covering mathematics, writing, social studies, science, and literature and arts. Except for the writing section, all the sections consist of multiple choice questions. The costs of acquiring a GED are relatively small. The pecuniary costs range from no cost in some states to around $50 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 (as measured by IQ) but exhibit more behavior and self-discipline problems and are less able to finish tasks. They switch jobs at a faster rate and are more likely to skip school, fight at school and work, use marijuana, 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; Heckman and Rubinstein, 2001).

Cavallo, Heckman, and Hsee (1998) and Heckman and Rubinstein (2001) have shown that when one controls for both cognitive and noncognitive abilities, there is no difference in earnings between a GED recipient and a dropout who has not acquired the certificate. Tyler, Murnane, and Willett (2000) obtain similar results except for young white dropouts who were in the margin of passing 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).

Because high school 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, because GED recipients do not earn higher wages than dropouts without the GED, the signal-earnings relation is not strictly monotone as in the traditional signalling models. As Heckman, Stixrud, and Urzua (2006) point out,

> Our evidence that multiple abilities determine schooling challenges the conventional single skill signalling model due to Arrow (1973) and Spence (1973). A special challenge is the GED program where the credential (the GED test) conveys multiple conflicting signals. GED recipients are smarter than other high school dropouts but they have lower noncognitive skills.

□ **The model.** In this subsection, we extend the basic framework to study the effect of the introduction of a pass-or-fail test such as the GED. We model the GED as a certifiable statement that only individuals with a sufficiently high combination of characteristics are able to reveal. Hence, we will add a disclosure dimension to the signalling model presented previously.[21]

We model the GED as an additional signal $h(\iota, \eta)$ that 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

[21] See, for example, Grossman (1981) and Okuno-Fujiwara, Postlewaite, and Suzumura (1990) for standard disclosure games.

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

where $\bar{g} \in \mathbb{R}_{++}$ 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.[22]

We assume that there is a nonnegative expected net benefit in passing the test. The positive expected net benefit may be a result of three characteristics of the GED. First, the costs of taking the exam are rather low. The median time studying for the GED is twenty hours and the monetary costs range from \$0 to \$50. Second, the GED provides recipients with the option of postsecondary schooling and joining the U.S. military. Third, there could be some nonmonetary benefits for being legally considered a high school graduate.

After obtaining the GED, the worker may choose whether or not to disclose this information to the employer. Hence, firms cannot distinguish between workers who have acquired a GED but chose not to disclose it from those who were unable to obtain the GED. We denote by $\tilde{h}$ the information disclosed by the worker. More specifically, $\tilde{h} = 1$ if $h = 1$ and the worker chooses to disclose this information. Otherwise, $\tilde{h} = 0$. Firms observe $\tilde{h}$ but not h. Therefore, employers now observe the amount of education y, the interview result g, and the GED $\tilde{h}$.

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

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

Because the GED exam is intensive in cognitive skills, we shall assume that the exam h emphasizes intelligence more than the interview g does.

Assumption 1. $\kappa > \alpha$.

Then, each worker with $\iota \geq \frac{\bar{g}-g}{\kappa-\alpha}$ would be able to pass the test. The graphs in Figure 5 separate the interval $[\iota_0, \iota_1]$ into three regions. The first graph depicts the case where $\frac{\bar{g}-g}{\kappa-\alpha} > \frac{g}{2\alpha}$, while the second graph represents the case where $\frac{\bar{g}-g}{\kappa-\alpha} < \frac{g}{2\alpha}$.

In the left region, workers have low intelligence so that education must be increasing in intelligence (CS_+ region) and the worker is unable to pass the test. On 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 $\frac{\bar{g}-g}{\kappa-\alpha} - \frac{g}{2\alpha}$. If $\frac{\bar{g}-g}{\kappa-\alpha} > \frac{g}{2\alpha}$ (first graph), some workers with types in the CS_- region are unable to receive $h = 1$. If $\frac{\bar{g}-g}{\kappa-\alpha} < \frac{g}{2\alpha}$ (second graph), some workers with types in the CS_+ region are able to pass the test.

The game consists of adding a last stage to the game presented in Section 2. Hence, the timing of the game is as follows. First, nature determines each worker's type according to density p. Second, workers choose their education $y^g(\iota, g)$ contingent on their types.[23] Subsequently, firms offer a wage $w^g(y, g, \tilde{h})$ conditional on observing (y, g) and on whether the worker will $(\tilde{h} = 1)$ or will not disclose the GED $(\tilde{h} = 0)$. Then, workers choose whether or not to acquire the GED and, if they do, whether or not to disclose this information.[24]

Consider the case where the firms' technology is intensive in noncognitive skills. Then, for any pool of workers, the one with higher perseverance/lower intelligence is the most productive.

[22] 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.

[23] We add the superscript g to differentiate the model where the GED is available from the one examined previously.

[24] In standard disclosure models, individuals first choose which information to reveal. A second stage consisting of a signalling game is easily introduced by assuming that the payoffs from the disclosure game are obtained by backward induction from the second stage. In the model presented above, however, the signalling game occurs in the first stage and the disclosure game occurs in the second stage. This is the natural assumption in our model because the choice of whether or not to take the GED is usually made after the worker has decided how much education to acquire.

FIGURE 5

WORKERS WHO ARE ABLE TO OBTAIN A GED

If firms could identify the most intelligent individuals in a pool of workers, they would offer them lower wages. Of course, a worker would never disclose the GED if this reduced her earnings, however. Hence, in the case where the firms' technology is intensive in noncognitive skills, allowing workers to take the GED does not affect education and wages (compared with the equilibrium obtained in Section 3). Thus, we say that, in this case, the GED is a neutral signal.[25] This conclusion, which is the main result of this section, is formally stated in the following proposition.

Proposition 5. Suppose that the firms' technology is intensive in noncognitive skills. Let $\{y(\iota, g), w(y(\iota, g), g)\}$ denote the quasi-separable equilibrium of the model where the GED is not available. There exists a quasi-separable equilibrium of the game where the GED is available such that acquiring the GED does not affect education and wages:

$$y^g\left(\iota, g, \tilde{h}^*(\iota, g)\right) = y(\iota, g),$$
$$w^g(y(\iota, g), g, \tilde{h}^*(\iota, g)) = w\left(y(\iota, g), g\right),$$
$$\tilde{h}^*(\iota, g) = 0 \quad \forall \iota \in [\iota_0, \gamma(\iota_0)],$$
$$h^*(\iota, g) = 1 \iff \iota \geq \frac{\bar{g} - g}{\kappa - \alpha}. \tag{16}$$

Proof. The result is trivial for a separating set. Assume two workers with $\iota \leq \frac{\bar{g}-g}{\kappa-\alpha} \leq \hat{\iota}$ are pooled in the same contract (otherwise, the signal is not informational). As acquiring the GED has positive net benefits and a worker can always choose not to disclose that she has obtained the GED, $\hat{\iota}$ will have $h = 1$ and ι will have $h = 0$. Suppose that type $\hat{\iota}$ chooses to disclose this information, that is, $\tilde{h}(\hat{\iota}, g) = 1$. Then, from Lemma 3, the firm would offer a lower salary for the type-$\hat{\iota}$ worker and a higher salary for ι. This cannot be an equilibrium, however, as the type-$\hat{\iota}$ worker's strategy is not optimal (condition i from Definition 1). Thus, all types who were discretely pooled in the quasi-separable equilibrium of the game where the GED was not available choose $\tilde{h} = 0$. Then, it follows that the separating and discrete pooling intervals as well as the conditions for the PBE obtained in Proposition 3 are the same in two the games, which concludes the proof.

Notice that, consistent with Heckman and Rubinstein (2001) and Cavallo, Heckman, and Hsee (1998), workers who have a GED have higher cognitive skills and lower noncognitive skills but receive the same wages as those who do not have it. However, as the result above holds for all $\bar{g} \in \mathbb{R}_{++}$, it follows that, unlike Cavallo, Heckman, and Hsee (1998) suggest, an increase in the GED standards, $\bar{g}$, would not affect the equilibrium education and wage schedules.[26] Furthermore, because the introduction of the GED does not affect the equilibrium profile of education, our model does not support the claim that, when the GED is neutral, it may discourage education (e.g., Cavallo, Heckman, and Hsee, 1998).

[25] When ι and $\gamma(\iota)$ are both able to obtain the GED (i.e., they are greater than $\frac{\bar{g}-g}{\kappa-\alpha}$), there also exist equilibria such that $\tilde{h}^*(\iota, g) = \tilde{h}^*(\gamma(\iota), g) = 1$. Because the GED does not disclose any information in this case, it does not affect education or wages.

[26] This implication of the model could be tested, as passing standards vary by states and have changed over time. Thus, one could test whether the neutrality of the GED is robust to different states and different periods of time.

A key assumption for the neutrality of the GED is that the firms' technology is intensive in noncognitive abilities.[27] Consider now the case where the firms' technology is intensive in cognitive skills. Then, for any pool of workers, the one who is able to acquire the GED is the most productive. Hence, by disclosing that one has a GED, a worker is able to obtain higher wages at no cost so that the GED is no longer neutral. The next proposition formally proves this result.

Proposition 6. Suppose that the firms' technology is intensive in cognitive skills and $\frac{\bar{g}-g}{\kappa-\alpha} \in [\iota_0, \gamma(\iota_0)]$. Then, there exists a quasi-separable equilibrium where the GED is nonneutral: for any $\iota \leq \frac{\bar{g}-g}{\kappa-\alpha} \leq \gamma(\iota)$,

$$w^g\left(y(\iota, g), g, \tilde{h}^*(\iota, g)\right) < w\left(y(\iota, g), g\right) < w\left(y\left(\gamma\left(\iota\right), g\right), g, \tilde{h}^*\left(\gamma\left(\iota\right), g\right)\right),$$

$$\tilde{h}^*(\iota, g) = 0, \quad \tilde{h}^*\left(\gamma\left(\iota\right), g\right) = 1,$$

$$h^*(\iota, g) = 1 \iff \iota \geq \frac{\bar{g}-g}{\kappa-\alpha}.$$

Proof. In this case, condition ii of Definition 1 implies that $w^g\ (y(\gamma\ (\iota), g), g, 1) = s(\gamma(\iota)) > s(\iota) = w^g\ (y(\iota, g), g, 0)$, where the inequality follows from Lemma 3. Then, type $\gamma\ (\iota)$ prefers to disclose $\tilde{h} = 1$ and type ι cannot pool with her because she cannot acquire the GED. The existence part follows the same steps as Section 3, except that now there will be separability in the two extremes.

Corollary 2. Suppose that the firms' technology is intensive in noncognitive skills. There exists a quasi-separable equilibrium such that an exam h that places more weight on noncognitive skills ($\kappa < \alpha$) is nonneutral.

A way to make the GED exam a nonneutral signal would be to put more emphasis on noncognitive 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 noncognitive skills, psychologists have developed tests that measure such skills, and they have been used by companies to screen workers (e.g., Sternberg, 1985).

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

6. Other applications

■ In the previous sections, we have shown that a two-dimensional signalling game featuring an additional exogenous signal can be reduced to a one-dimensional signalling game where the SCP is violated. Moreover, we have characterized the equilibrium of this one-dimensional signalling game.

Although the article is presented in a job-market environment, it can be employed in a wide variety of environments. In this section, we briefly discuss some examples.

First, consider a corporate finance context, where firms may use dividends in order to signal future earnings. Reinterpret ι as current earnings, η as future earnings, y as the amount of dividends

[27] The neutrality of the GED does not rely on the assumption that education does not affect the ability to pass the GED exam. For example, suppose that an individual would be able to pass the GED if $\kappa\iota + \eta + \beta y \geq \bar{g}$. Then, the shaded area in Figure 5 would depend on y, but if two workers were discretely pooled in a contract, the one who could pass the test would still be the least productive worker.

paid, and g as representing a specialist's classification of the profitability of the firm. Then, if the firm's and the specialist's time preferences were not aligned, the SCP would not hold in general (see Appendix A). This misalignment might be a result of credit constraints. It could also be a consequence of an (unmodelled) CEO remuneration contract that induces greater short-term orientation.[28] Consistently with Benartzi, Michaely, and Thaler (1997), our model would then predict a nonmonotonic relation between dividends and future earnings. Furthermore, higher dividends would be a mixed signal. Depending on the time preference of the firm and investors, it could signal both high present earnings and low future earnings or low present earnings and high future earnings.[29]

In an international finance context, we could reinterpret ι as the government's commitment to maintaining a fixed exchange rate, η as the quality of the fundamentals of the country, and y as the interest rate. The signal g could denote the country's risk classification or some other indicator of its fundamentals. As long as the risk classification is not perfectly aligned with the government's preferences in the sense that their marginal rates of substitution between ι and η cannot be ordered, the SCP would not hold. In that case, consistently with the evidence from Drazen and Hubrich (2003), our model would predict that interest rates are mixed signals, as they indicate that the government is committed to maintaining a fixed exchange rate, but may also signal weak fundamentals. Furthermore, countersignalling implies that it could be optimal for a country to choose lower interest rates in order to signal strong fundamentals.

Another application is a model where firms choose the amount of advertising expenditures in order to signal product quality, captured by a two-dimensional type (ι, η). Reinterpreting y as advertising expenditures, g as additional information obtained from other sources (e.g., word-of-mouth advertising, magazine reviews, etc.), one can easily apply the model presented above. In that model, advertising may be a mixed signal. Furthermore, consistent with the evidence from Caves and Greene (1996), Clements (2004), and Orzach, Overgaard, and Tauman (2002), there would not be an increasing relationship between advertising and product quality.

7. Conclusion

■ In this article, we presented a model of mixed signals. We demonstrated that when firms have access to an interview technology, the two-dimensional model can be reduced to a one-dimensional model where the single-crossing property may not hold. When this is the case, the equilibrium features countersignalling in the sense that signals are nonmonotone in the worker's productivity.

It was shown that countersignalling occurs if, and only if, the schooling technology differs from the firm's technology. Moreover, the countersignalling 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 a more diverse combination of skills from those required in the schooling process.

We have extended the basic model in order to analyze the GED exam. It was shown that, consistently with the empirical evidence, a GED recipient has above-average cognitive skills and below-average noncognitive skills. When cognitive skills are more valued in the labor market, this new information affects equilibrium wages. However, when noncognitive 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 firm's main concern is usually about the worker's noncognitive skills, a nonneutral signal should assign more weight to this kind of skill. Thus, changing its focus to noncognitive skills would turn the GED into

[28] Bolton, Scheinkman, and Xiong (2005) have shown that when the equilibrium stock prices may differ from the fundamental value, the optimal contract may induce greater short-term orientation.

[29] See Araujo, Moreira, and Tsuchida (2004) for a model of dividends as signals where the SCP does not hold.

a nonneutral 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.

Our results provide evidence of the importance of the failure of the single-crossing property in explaining 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 noncognitive ability receives the same wages as another with low cognitive ability and high noncognitive ability while an individual with intermediate abilities does not is evidence of lack of the single-crossing property.

This article also has a technical contribution, as it characterizes the equilibrium of a signalling model where the single-crossing condition does not hold. This framework can be employed in a wide variety of environments such as advertising, corporate finance, and international finance.

Appendix A

□ **Robustness of the single-crossing property.** In this section, we characterize the set of functions c and g that satisfy the single-crossing property (SCP). We will show that the SCP does not hold as long as the interview technology and the schooling technology cannot be ordered according to their technical rates of substitution.

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

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

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

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

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

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

if and only if $g(\iota, \eta) = \bar{g}$.

Moreover,

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

Substituting into the cost function, it follows that

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

Hence,

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

Thus, the SCP holds if, and only if, $\frac{r_\iota}{r_\eta} - \frac{g_\iota}{g_\eta}$ 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 interview technology can be ordered.

Suppose, for example, that w and g are both constant elasticity of substitution (CES) functions[30]:

$$r = [\alpha_1 \iota^\rho + \alpha_2 \eta^\rho]^{\frac{1}{\rho}},$$

$$g = [\beta_1 \iota^\gamma + \beta_2 \eta^\gamma]^{\frac{1}{\gamma}}.$$

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

When the SCP does not hold and the CS_+ and CS_- regions are not independent of y, the necessary conditions may no longer be sufficient and we cannot guarantee that an equilibrium exists. If it exists, however, some results presented in this article still hold. As in Remark 1, it can be shown that we cannot in general have a fully separating equilibrium when the SCP is violated. The equilibrium must be such that y is increasing in the CS_+ region and decreasing in the CS_- region so that, as long as it does not feature complete pooling, y will not be monotone.

In a previous version of the article, we have analyzed the case where education affects the interview result so that the CS_+ and CS_- regions are no longer independent of y. We have shown that if a quasi-separable equilibrium exists, then our results on countersignalling and the GED still hold.[31]

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

[31] See Araujo, Gottlieb, and Moreira (2007).

Appendix B

■ Proofs of Lemmas 1–7 and Propositions 1–3 following:

Proof of Lemma 1. Because the first claim of this lemma is a particular case of Lemma 4, we will only prove Lemma 4. The second claim follows from equation (9).

Proof of Lemma 2. If $\{w(y(\iota, g)), y(\iota, g)\}$ is an incentive-compatible profile of education and wages, it must satisfy

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

The first-order condition of the program above yields

$$w_y\left(y(\iota, g), g\right) = c_y\left(\iota, g, y(\iota, g)\right). \tag{B1}$$

Suppose that $y(\iota, g) = y(\tilde{\iota}, g)$ for some regular types ι and $\tilde{\iota}$. Substituting in equation (B1) yields $c_y(\iota, g, y(\iota, g)) = c_y(\tilde{\iota}, g, y(\tilde{\iota}, g))$.

Proof of Lemma 3. Let $\iota > \hat{\iota}$ be two discretely pooled workers and notice that $\alpha\hat{\iota} = \eta$ and $\alpha\iota = \hat{\eta}$. Substituting in the firm's technology yields

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

Proof of Lemma 4. Define $U\left(\hat{\iota}, \iota\right)$ as the expected utility received by a type-(ι, g) individual who gets a contract designed for type $\left(\hat{\iota}, g\right)$:

$$U\left(\hat{\iota}, \iota\right) = P\left(\hat{\iota}, g\right) s\left(\hat{\iota}, g\right) + P\left(\frac{g}{\alpha} - \hat{\iota}, g\right) s\left(\frac{g}{\alpha} - \hat{\iota}, g\right) - c\left(\iota, g, y\left(\hat{\iota}, g\right)\right).$$

The incentive-compatibility constraint is

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

The local first-order condition is

$$U_{\hat{\iota}}\left(\hat{\iota}, \iota\right)|_{\hat{\iota}=\iota} = 0, \quad \forall \iota \in [\iota_0, \iota_1]. \tag{B2}$$

Calculating the derivative above yields

$$\begin{aligned} c_y\left(\iota, g, y(\iota, g)\right) y_\iota(\iota, g) = \; & P_\iota(\iota, g) s(\iota, g) + P(\iota, g) s_\iota(\iota, g) \\ & - P_\iota\left(\frac{g}{\alpha} - \iota, g\right) s\left(\frac{g}{\alpha} - \iota, g\right) - P\left(\frac{g}{\alpha} - \iota, g\right) s_\iota\left(\frac{g}{\alpha} - \iota, g\right). \end{aligned}$$

Because $s_\iota(x, g) = \frac{bg - \alpha x}{x(g - \alpha x)} s(x, g)$ and $c_y(\iota, g, y(\iota, g)) = \frac{1}{\iota(g - \alpha\iota)}$, the expression becomes

$$\begin{aligned} y_\iota(\iota, g) = \; & s(\iota, g)\left[P_\iota(\iota, g)\iota(g - \alpha\iota) + P(\iota, g)(bg - \alpha\iota)\right] \\ & + s\left(\frac{g}{\alpha} - \iota, g\right)\left[P\left(\frac{g}{\alpha} - \iota, g\right)\left[g(1 - b) - \alpha\iota\right] - P_\iota\left(\frac{g}{\alpha} - \iota, g\right)\iota(g - \alpha\iota)\right]. \end{aligned}$$

Using the fact that $P(\frac{g}{\alpha} - \iota, g) = 1 - P(\iota, g)$ for all ι, we obtain equation (13).

Differentiating equation (B2) with respect to ι, we obtain

$$U_{\hat{\iota}\hat{\iota}}\left(\iota, \iota\right) + U_{\hat{\iota}\iota}\left(\iota, \iota\right) = 0. \tag{B3}$$

The necessary second-order condition is

$$U_{\hat{\iota}\hat{\iota}}\left(\iota, \iota\right) \le 0. \tag{B4}$$

Substituting (B3) in (B4), it follows that

$$U_{\hat{\iota}\iota}\left(\iota, \iota\right) = -c_{y\iota}\left(\iota, g, y(\iota, g)\right) y_\iota(\iota, g) \ge 0.$$

Substituting $c_{y\iota}(\iota, g, y) = -\frac{g - 2\alpha\iota}{(\iota(g - \alpha\iota))^2}$ in the inequality above establishes (14).

Proof of Proposition 1. Suppose that wages are not strictly increasing in education. Then, there exist types ι and $\tilde{\iota}$ such that

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

This is not incentive compatible, however, as

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

concluding the first part of the proof.

In order to establish the concavity of w, consider the incentive-compatibility constraint

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

The second-order condition (necessary) is[32]

$$w_{yy}(y(\iota, g), g) \leq 0.$$

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

Proof of Lemma 5. Suppose that ι is an interior point of either a separating set or a discrete pooling set. Then, as y is continuous in ι (because it solves a differential equation) it follows that

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

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

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

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,

$$\lim_{x \to \iota_+} U(x, x) = s(\iota, g) - \frac{\lim_{x \to \iota_+} y(x, g)}{\iota(g - \alpha\iota)},$$

$$\lim_{x \to \iota_-} U(x, x) = P(\iota, g)s(\iota, g) + [1 - P(\iota, g)]\, s(\gamma(\iota, g), g) - \frac{y(\iota, g)}{\iota(g - \alpha\iota)}$$

because $y(\cdot, g)$ is right continuous at ι.

Thus, the inequality can be written as

$$y(\iota, g) \geq \lim_{x \to \iota_+} y(x, g) + \iota(g - \alpha\iota)[1 - P(\iota, g)][s(\gamma(\iota, g), g) - s(\iota, g)].$$

Another necessary condition for incentive compatibility is

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

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

Using the definition of the indirect utility, we get

$$y(\iota, g) \leq \lim_{x \to \iota_+} y(x, g) + \iota(g - \alpha\iota)[1 - P(\iota, g)][s(\gamma(\iota, g), g) - s(\iota, g)].$$

Combining these two necessary conditions, we obtain equation (15).

Substituting in the indirect utility function, it follows that $U(\iota, \iota) = \lim_{x \to \iota} U(x, \iota)$.

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 (as there is no continuous pooling in a quasi-separable equilibrium). Assume that some type in $[\iota_0, \gamma(\iota_0)]$ is separated. Then, there must be a $\iota \in [\iota_0, \frac{g}{2\alpha}]$ such that $[\iota, \frac{g}{2\alpha}]$ is a discrete pooling set and $[\iota - \varepsilon, \iota]$ is a separated set for some $\varepsilon > 0$. From equation (15), it follows that y jumps upward when the types become separated. This is not incentive compatible, however, 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 wage must be $s(\iota_1, g)$ in any separating equilibrium (which is the lowest wage as ι_1 is the least productive type). However, she would receive a wage of at least $s(\iota_1, g)$ if she chose $y = 0$. As $y = 0$ implies a lower signalling cost and does not reduce her utility, she would be strictly better off by doing so.

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

Proof of Proposition 3. Let where $y(\iota, g)$ be given by the solution to the differential equations from Lemma 1 and Lemma 4 with the boundary conditions from Lemma 5 and Lemma 7. Define $w(y, g)$ as in condition ii from Definition 1. Let μ be a Dirac measure concentrated at ι in the interval $[\gamma(\iota_0, g], \iota_1]$ and $P(\iota, g)$ in the interval $[\iota_0, \gamma(\iota_0, g)]$ for y in the range of signals.[33]

By construction, in order to show that $\{y(\iota, g), w(y, g)\}$ and $\mu(.|y, g)$ is a PBE, it suffices to establish that

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

for all $\iota \in [\iota_0, \iota_1]$.

First, observe that the first-order condition of the program above is equivalent to equation (7) for $\iota \in [\gamma(\iota_0, g], \iota_1]$ and equation (13) for $\iota \in [\iota_0, \gamma(\iota_0, g)$ and, therefore, are satisfied by $y(\iota, g)$. Moreover, the (global) second-order condition is equivalent to

$$w_{yy}(y, g) - c_{yy}(\iota, g, y) \leq 0$$

for y in the range of signals.

From equation (1), $c_{yy}(\iota, g, y) = 0$. Then, because $w_y(y, g) = c_y(\iota, g, y)$ for $y = y(\iota, g)$ by the first-order condition,

$$w_{yy}(y, g)y_\iota = c_{yy}(\iota, g, y)y_\iota + c_{y\iota}(\iota, g, y)$$

for $y = y(\iota, g)$. Thus, whenever $y_\iota \neq 0$,

$$w_{yy}(y, g) = \frac{c_{y\iota}(\iota, g, y)}{y_\iota(\iota, g)} \leq 0$$

for $y = y(\iota, g)$, by equations (8) and (14). Therefore, the (global) second-order condition holds.

To complete the first part of the proof, we have to show that this PBE is the quasi-separable equilibrium. This is clear, however, by inspection of Definition 3 and Lemma 7. Existence and uniqueness of the quasi-separable equilibrium follow from the fact that both differential equations are Lipschitz.

References

ARAUJO, A. AND MOREIRA, H. "Adverse Selection Problems without the Spence-Mirrlees Condition." Working Paper no. 425, Getulio Vargas Foundation, 2001.

——— AND TSUCHIDA, M. "Do Dividends Signal More Earnings?" Working Paper no. 524, Getulio Vargas Foundation, 2004.

ARAUJO, A., GOTTLIEB, D., AND MOREIRA, H. "A Model of Mixed Signals with Applications to Countersignaling and the GED." Working Paper, Getulio Vargas Foundation, 2007.

ARROW, K.J. "Higher Education as a Filter." *Journal of Public Economics,* Vol. 2 (1973), pp. 193–216.

BARRICK, M. AND MOUNT, M. "The Big Five Personality Dimensions and Job Performance: A Meta-Analysis." *Personnel Psychology,* Vol. 44 (1991), pp. 1–26.

BENABOU, R. AND TIROLE, J. "Incentives and Prosocial Behavior." *American Economic Review,* Vol. 96 (2006), pp. 1652–1678.

BENARTZI, S., MICHAELY, R., AND THALER, R.H. "Do Changes in Dividends Signal the Future or the Past?" *Journal of Finance,* Vol. 52 (1997), pp. 1007–1034.

BOESEL, D., ALSALAM, N., AND SMITH, T.M. *Educational and Labor Market Performance of GED Recipients*. Washington, DC: National Library of Education, Office of Educational Research and Improvement, U.S. Department of Education, 1998.

BOLTON, P., SCHEINKMAN, J.A., AND XIONG, W. "Executive Compensation and Short-termist Behavior in Speculative Markets." *Review of Economic Studies,* Vol. 73 (2005), pp. 577–610.

BOWLES, S. AND GINTIS, H. "Schooling in Capitalist America Revisited." *Sociology of Education*, Vol. 75 (2001), pp. 1–18.

CAMERON, S. AND HECKMAN, J. "The Nonequivalence of High School Equivalents." *Journal of Labor Economics,* Vol. 11 (1993), pp. 1–47.

CARNEIRO, P. AND HECKMAN, J. "Human Capital Policy." Working Paper no. 9495, National Bureau of Economic Research, 2003.

CAVALLO, A., HECKMAN, J., AND HSEE, J. "The GED Is a Mixed Signal." Unpublished manuscript presented at the meetings of the American Economic Association, New York, January 3, 1998.

CAVES, R.E. AND GREENE, D.P. "Brands' Quality Levels, Prices, and Advertising Outlays: Empirical Evidence on Signals and Information Costs." *International Journal of Industrial Organization*, Vol. 14 (1996), pp. 29–52.

[33] For y outside the range of signals, let μ be a Dirac measure concentrated at ι_1 (which is the least productive type by Lemma 3).

CAWLEY, J., CONNEELY, K., HECKMAN, J., AND VYTLACIL, E. "Measuring the Effects of Cognitive Ability." Working Paper no. 5645, National Bureau of Economic Research, 1996.

CLEMENTS, M.T. "Low Quality as a Signal of High Quality." Working Paper, University of Texas at Austin, 2004.

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

DRAZEN, A. AND HUBRICH, S. "Mixed Signals in Defending the Exchange Rate: What Do the Data Say?" CEPR Discussion Paper no. 4050, 2003.

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

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

FELTOVICH, N., HARBAUGH, R., AND TO, T. "Too Cool for School? Signaling and Countersignaling." *RAND Journal of Economics,* Vol. 33 (2002), pp. 630–649.

GREEN, J.R. AND LAFFONT, J.J. "Competition on Many Fronts: A Stackelberg Signalling Equilibrium." *Games and Economic Behavior,* Vol. 2 (1990), pp. 247–272.

GROSSMAN, S. "The Informational Role of Warranties and Private Disclosure about Product Quality." *Journal of Law and Economics,* Vol. 24 (1981), pp. 461–483.

HECKMAN, J. AND RUBINSTEIN, Y. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *AEA Papers and Proceedings,* Vol. 91 (2001), pp. 145–149.

———, STIXRUD, J., AND URZUA, S. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes." Working Paper No. W12006, National Bureau of Economic Research, 2006.

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

HVIDE, H. "Education and the Allocation of Talent." *Journal of Labor Economics,* Vol. 21 (2003), pp. 945–976.

KHOLLEPPEL, L. "Multidimensional 'Market Signaling'." Discussion Paper, Universität Bonn, 1983.

KLEIN, R., SPADY, R., AND WEISS, A. "Factors Affecting the Output and Quit Propensities of Production Workers." *Review of Economic Studies,* Vol. 58 (1991), pp. 929–954.

OKUNO-FUJIWARA, M., POSTLEWAITE, A., AND SUZUMURA, K. "Strategic Information Revelation." *Review of Economic Studies,* Vol. 57 (1990), pp. 25–47.

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

ORZACH, R., OVERGAARD, P.B., AND TAUMAN, Y. "Modest Advertising Signals Strength." *RAND Journal of Economics,* Vol. 33 (2002), pp. 340–358.

PINKER, S. *How the Mind Works*. New York: Norton, 1999.

QUINZII, M. AND ROCHET, J.-C. "Multidimensional Signaling." *Journal of Mathematical Economics,* Vol. 14 (1985), pp. 261–284.

RILEY, J. "Informational Equilibrium." *Econometrica,* Vol. 47 (1979), pp. 331–359.

SPENCE, M. *Market Signalling: Information Transfer in Hiring and Related Processes*. Cambridge, Mass.: Harvard University Press, 1973.

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

TYLER, J., MURNANE, R., AND WILLETT, J. "Estimating the Labor Market Signaling Value of the GED." *Quarterly Journal of Economics,* Vol. 115 (2000), pp. 431–468.