Market signaling with grades

Brendan Daley, Brett Green

Journal of Economic Theory 151 (2014) 114–145

- The typical signaling model assumes that only observable, costly actions can convey information about the sender.
- These models implicitly overlook the existence of grades.
  - A grade refers to any imperfect public message about the sender's type (e.g., test scores, analyst ratings, product reviews).
- In this paper, we study how strategic agents behave when both channels are available for information transmission and consider an amended version of Spence's job-market signaling model.

1 Introduction

- Our first insight is that, in the presence of an informative test, some degree of pooling on the education level is more plausible than the widely adopted prediction of separation.
- Our second key insight is that the addition of grades can resolve the discontinuity of the equilibrium prediction as the prior converges to degeneracy.
- The presence of an RC-Informative test also has implications for Pareto efficiency.
- Focusing on the equilibrium satisfying D1, we explore the welfare implications of grades.
- We generalize our main results by extending our analysis to a more general class of preferences and by allowing the accuracy of the test to vary with the action.

- In a noisy signaling model, receivers do not perfectly observe the sender's action, but rather a noisy signal whose distribution depends on it ((Matthews and Mirman (1983), Carlsson and Dasgupta (1997)).
- The model with grades is a more accurate description of the strategic situation: employers observe both ***years of schooling*** and ***grades*** on transcripts.
- Our work illustrates the trade-off between these two channels and explains the importance of both in delivering our predictions.

- Ours is not the first paper to introduce additional information in a signaling framework (Weiss, 1983; Fang, 2001; Angeletos and Pavan , 2006).
- Feltovich et al. (2002) identify conditions for “countersignaling” equilibria to exist.
- Alos-Ferrer and Prat (2012) amend the canonical two-type job-market signaling model to one after hiring, and show that the sender’s type is gradually revealed via on-the-job employer learning.

- In contrast to these two papers, we fully characterize the set of equilibria in a two-type model and show that the main economic insights extend to a model with more types.
- We also consider a more general structure for the external information, including allowing the quality of information to vary with the sender’s costly action.
- Finally, on the more technical side, we show that the double-crossing property (Matthews and Moore (1987)) arises naturally in our model for the relevant indifference curves and facilitates a tractable equilibrium analysis.

- With in the canonical signaling example of Spence(1973).

- First, the worker bases on its type and strategy to take an action(education) $x \in S_t$.
- After the sender chooses an action, the grade $g \in \mathbb{R}$ is then realized.
- Next, each receiver i simultaneously makes an offer $W_i(x, g)$ to the sender based on observing x and g.
  - If the sender accepts offer:

    $U^S = W - C_t x$ where $0 < C_H < C_L$

    $U^R = V_t - W_i$ where $0 < V_L < V_H$
  - If the sender rejects offer:

    $U^S = -C_t x$

    $U^R = 0$

3.1 Grades and tests

- We refer to the pair of probability density functions $\{f_L, f_H\}$ as a test.
- Let $R(g) \equiv \frac{f_L(g)}{f_H(g)}$, measure the informativeness of grade g.
- Since we focus on statistically informative tests, we require the test to satisfy the following conditions:
  - T.1 For both types t, f_t is continuous almost everywhere.
  - T.2 Grades are boundedly-informative: $inf_g\ R(g) > 0$ and $sup_g\ R(g) < \infty$.
  - T.3 The Monotone Likelihood Ratio Property (MLRP) holds with R weakly decreasing over the common support of f_L and f_H .

- We are also interested in measuring the informativeness of tests.
- Based on Blackwell (1950) and Lehmann (1988), the crucial notion of informativeness in our analysis is the low type's expected likelihood ratio, $E[R(g)|t = L]$.
  - The higher is $E[R(g)|t = L]$, the more informative the test.
- Intuitively, what matters for equilibrium analysis is not just the test informativeness, but rather the test informativeness relative to the high type's advantage in taking the costly action.

Definition 3.3

The test is RC-Informative if and only if $E[R(g)|L] > \frac{c_L}{c_H}$.

Definition 3.3

The test is RC-Informative if and only if $E[R(g)|L] > \frac{c_L}{c_H}$.

- Example 3.4

A symmetric binary test is RC-Informative if and only if:

$$p > \frac{1}{2}\left(1 + \frac{\sqrt{(\frac{C_L}{C_H})^2 + 2\frac{C_L}{C_H} - 3}}{\frac{C_L}{C_H} + 3}\right)$$

where the right-hand side is increasing in the cost advantage.

For instance, if the cost advantage is $\frac{c_L}{c_H} = \frac{3}{2}$, then the test is RC-Informative if and only if $p > \frac{2}{3}$.

Remark 3.5

Our assumption that $C_L > C_H$ corresponds to the standard single-crossing property, facilitating comparison to the standard gradeless model. However, our analysis does not rely on this assumption. When $C_L = C_H$, any statistically informative test is RC-Informative, and all of our results under RC-Informativeness hold in this environment.

- We use perfect Bayesian equilibrium (PBE) as our solution concept.
- After observing x and g, receivers update to some final belief $\mu_f(x,g) \equiv \Pr(t = H|xg)$.
    - $\mu_f(x,g) = \frac{\mu(x)}{\mu(x)+(1-\mu(x))R(g)}$
- Let $W(x,g)$ be the highest offer from the receivers after observing x and g.
    - $W(x,g)=\mu_f(x,g)V_H + (1-\mu_f(x,g))V_L$
- Type t's (highest) expected offer:
    - $w_t(\mu)=\int \frac{\mu V_H+(1-\mu)V_L}{\mu+(1-\mu)r(g)} f_t(g)dg$
- The expected utility of a type- t worker depends only on his chosen action and the resultant interim belief:
    - $u_t(x,\mu)=w_t(\mu)-C_t x$
    - $x_0 \in argmax_x u_t(x,\mu(x))$

- **Without grades**, the indifference curves of interest are those over the space of actions and offers (or, equivalently, **actions and final beliefs**).
- **With grades**, it is indifference curves over the space of **actions and interim beliefs** that are crucial for analysis. We refer to these as Belief Indifference Curves (BICs).

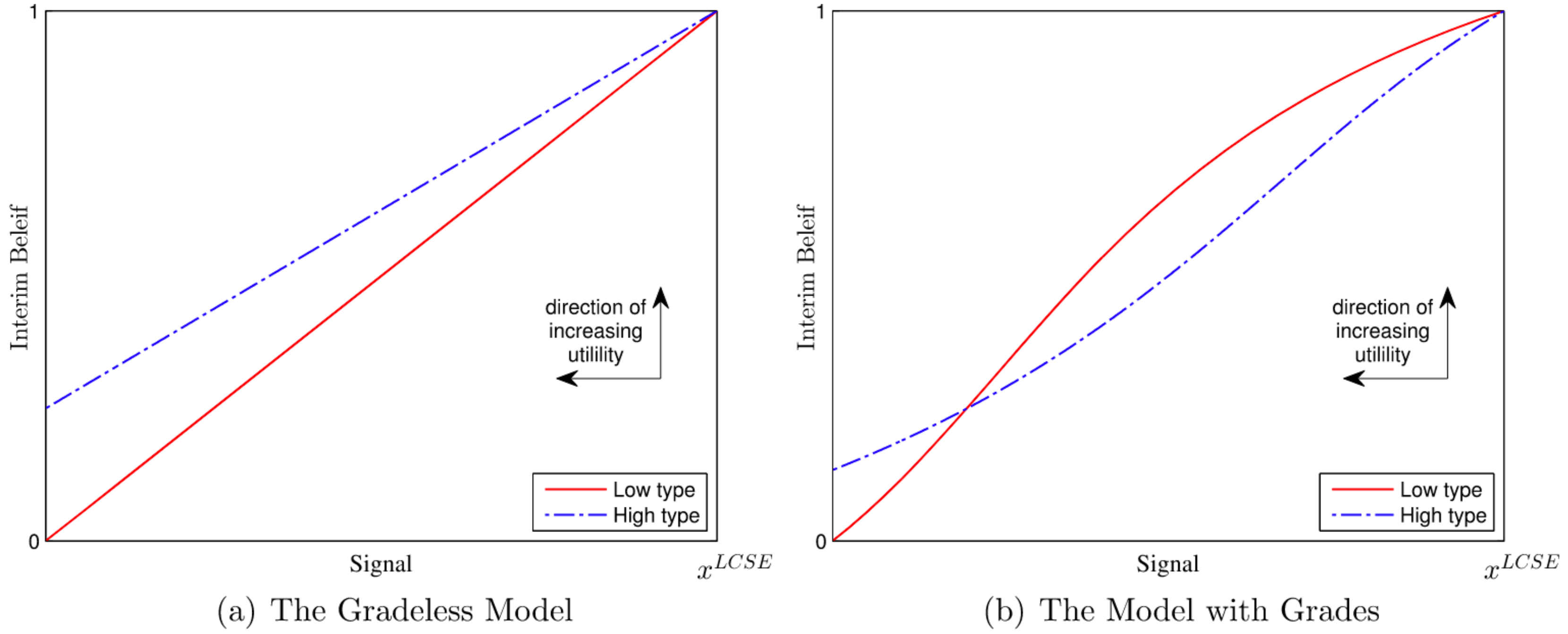


Fig. 1. BICs for the LCSE payoffs.

- BICs have a useful feature, termed the monotone relative slope property (MRSP) .
- Specifically, there exists a unique $\mu^* \in (0,1]$ which changes the relative slope of the two types.
  - Notice that BICs satisfy the single-crossing property if and only if $\mu^* = 1$.
- The high type has two advantages: a cost advantage and a grade advantage.
  - The cost advantage is independent of the interim belief and can be measured by $\frac{c_L}{c_H}$;
  - The grade advantage depends on receivers' interim belief.

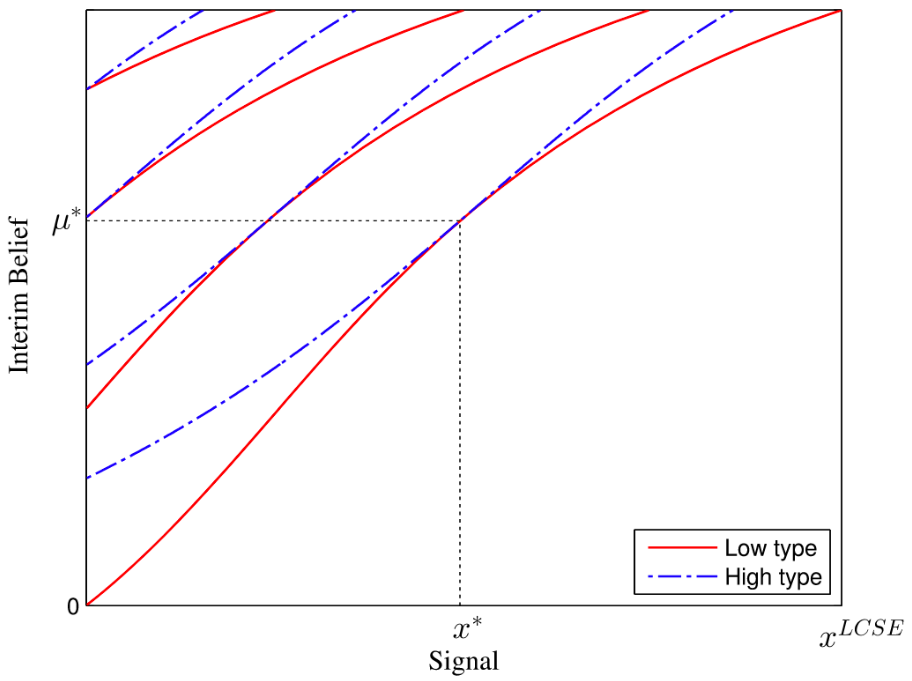


Fig. 2. Illustration of the monotone relative slope property.

- When μ is close to zero, an increase in μ makes the test more important and offers are more sensitive to grades. This benefits high type and the grade advantage is increasing in this region.
- In contrast, for sufficiently large μ, further increases in μ benefit the low type.
- Based on this relationship between interim belief and test importance, MRSP results from the relative slopes of the BICs.

Lemma 3.6

The following statements are equivalent.
1. The test is RC-Informative.
2. At $\mu = 1$, the high type's BIC is steeper than the low type's BIC.
3. $\mu^* < 1$.

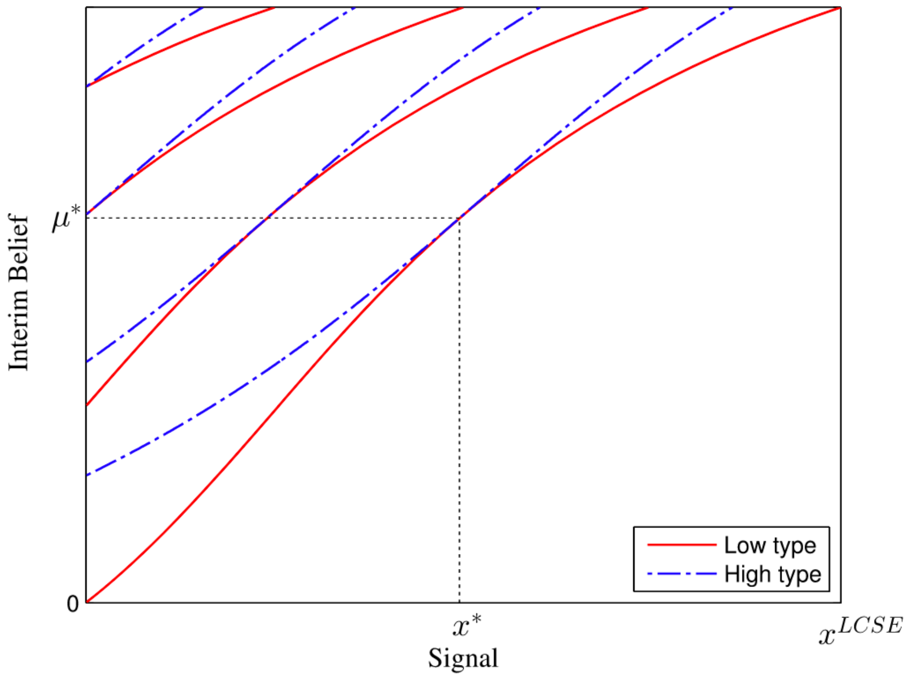


Fig. 2. Illustration of the monotone relative slope property.

- There exists a new form of equilibria we designate common support equilibria.
- In a common support equilibrium $S_L = S_H$, but $\Upsilon_L \neq \Upsilon_H$. The characteristic is no on-path action can identify the sender's type perfectly, and multiple actions are on the equilibrium path.
- MRSP implies that in any common support equilibrium $S_L = S_H = \{x_1, x_2\}$, where $x_1 < x_2$ and $\mu(x_1) < \mu_0 < \mu(x_2)$

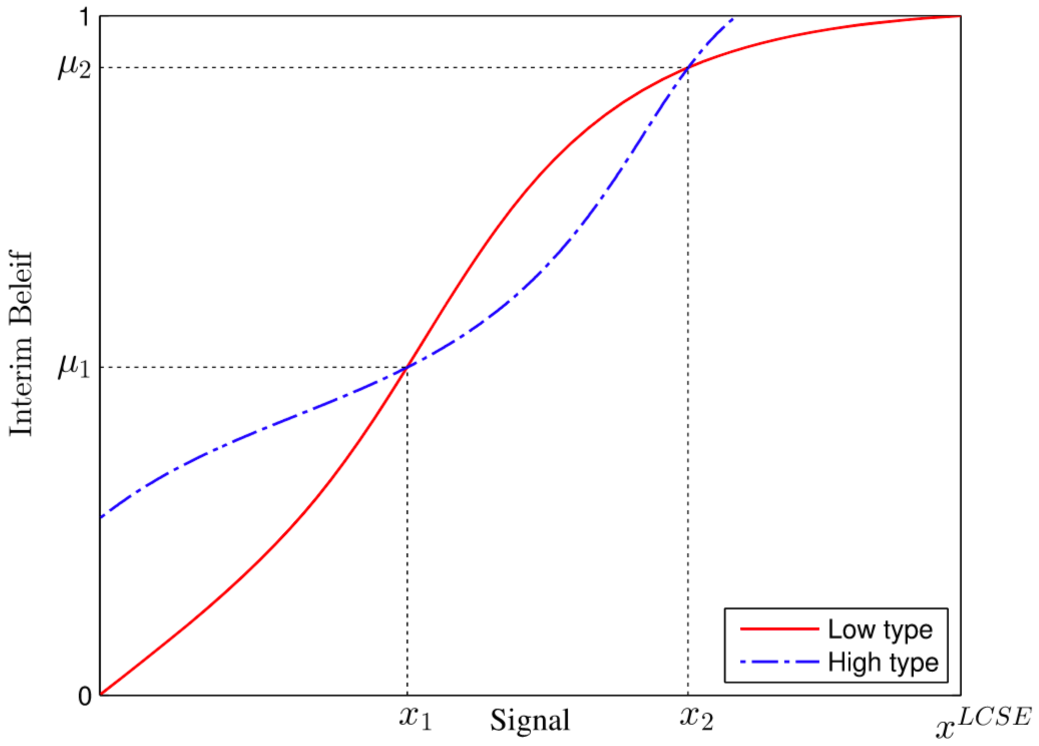
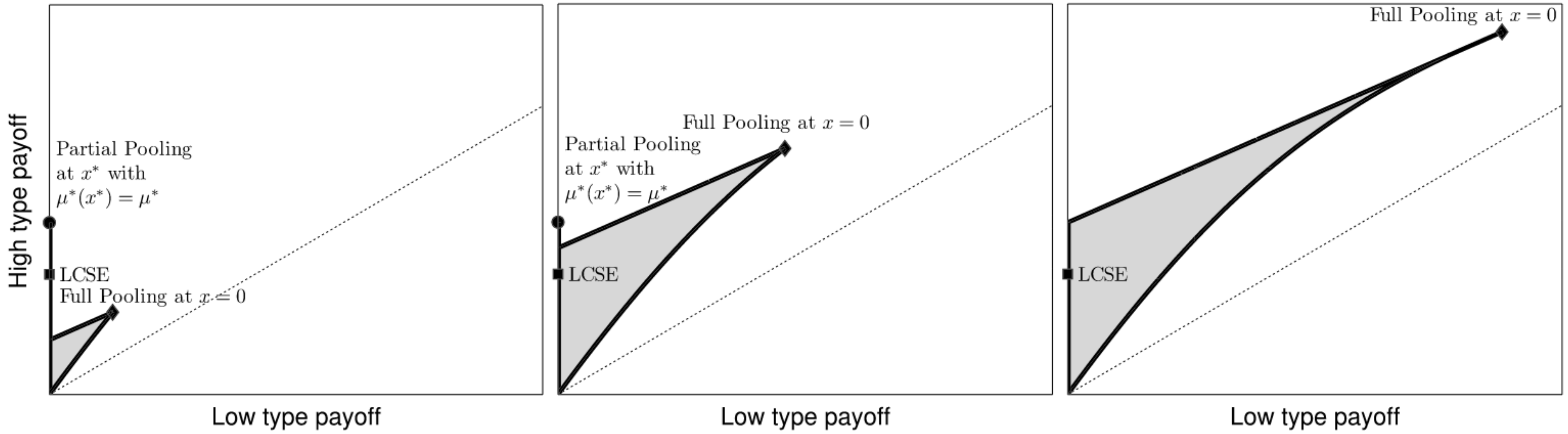


Fig. 3. BICs for a common support equilibrium.

Proposition 3.7

For any μ_0, if the test is RC-Informative, then grades convey payoff-relevant information in every Pareto efficient equilibrium. If the test is not RC-Informative, then the LCSE is a Pareto efficient equilibrium if and only if $\mu_0 \le \underline{\mu}$.



(a) For $\mu_0 < \underline{\mu}$ (b) For $\mu_0 \in (\underline{\mu}, \mu^*)$ (c) For $\mu_0 > \mu^*$

Proposition 3.8

If the test is RC-Informative, then:

- If $\mu_0 > \mu^*$, the unique D1 equilibrium is full pooling at $\mathrm{x} = 0$.
- If $\mu_0 < \mu^*$, the unique D1 equilibrium is partial pooling. The high type chooses $x^*$with probability 1, and the low type mixes over $x = 0$ and x^* such that $\mu(x^*) = \mu^*$.
- If $\mu_0 < \mu^*$, all D1 equilibria are full pooling and can be supported at x iff $x \in [0, x^*]$.

If the test is not RC-Informative, then for all priors the unique D1 equilibrium is the LCSE.

3.5 Equilibrium under D1: Senders rely on informative tests

- To see why some degree of pooling must occur when the test is RC-Informative, consider the LCSE and suppose receivers observe an off-path action of $x^{LCSE} - \epsilon$.
  - Clearly, $\mu(x^{LCSE} - \epsilon)$ must be close to one in order for this to be a profitable deviation for either type.
  - MRSP implies the high type's BIC is steeper at all $\mu > \mu^*$.
- In accordance with D1, receivers will assign probability one to $t = H$ after observing the deviation, making it profitable for both types, and breaking the candidate separating equilibrium.

- In the D1 equilibrium, the high type uses the costly action to influence the receivers' interim belief only as long as he has a relative advantage in doing so.
- When the test is RC-Informative, the high type does not fully separate through the costly action.
  - When the prior is below μ^*, he uses the costly action to the point where the interim belief reaches μ^*;
  - When the prior is above μ^*, any attempt to influence the interim belief is viewed as an attempt to de-emphasize the test, and hence both types fully pool on the costly action at x = 0.

3.6 Equilibrium convergence

- We introduce the following two notions of convergence.

Definition 3.9

Let $\{\mu_0^n\}$ be a sequence of priors converging to μ_0, and $\{\gamma_L^n, \gamma_H^n\}$ be any sequence of strategy profiles such that γ_L^n, γ_H^n is an equilibrium when the prior is μ_0^n.

- The set of equilibrium strategy profiles *converges type-by-type* to a distribution γ if for every sequence $\{\gamma_L^n, \gamma_H^n\}$, γ_t^n converges in distribution to γ for all t .
- The set of equilibrium strategy profiles converges in total mass to a distribution γ if for every sequence $\{\gamma_L^n, \gamma_H^n\}$, $\mu_0^n \cdot \gamma_H^n + (1 - \mu_0^n) \cdot \gamma_L^n$ converges in distribution to γ.

Corollary 3.10

As $\mu_0 \to 1$, by either notion of convergence (type-by-type or in total mass), the D1 equilibrium converges to the complete-information outcome if and only if the test is RC-Informative.

Reliance on the test.

As the informativeness of the test increases, the high type relies more on it, and less on the costly action. Analogously, as the high type's cost advantage increases, he relies more on the costly action, and less on the test.

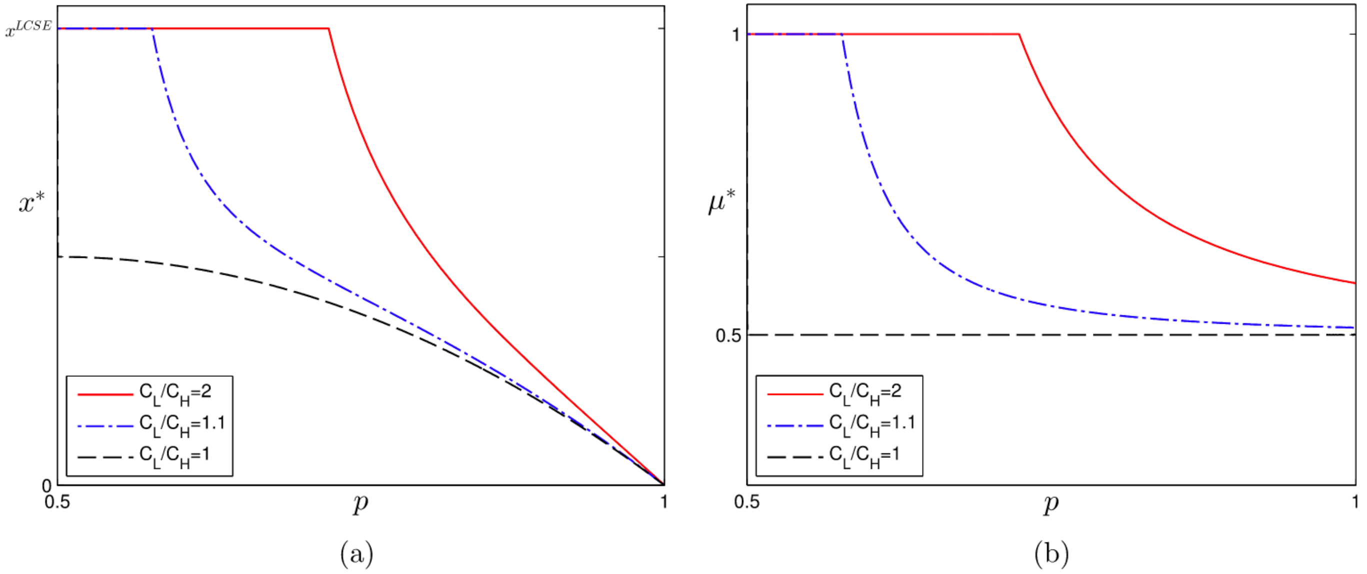
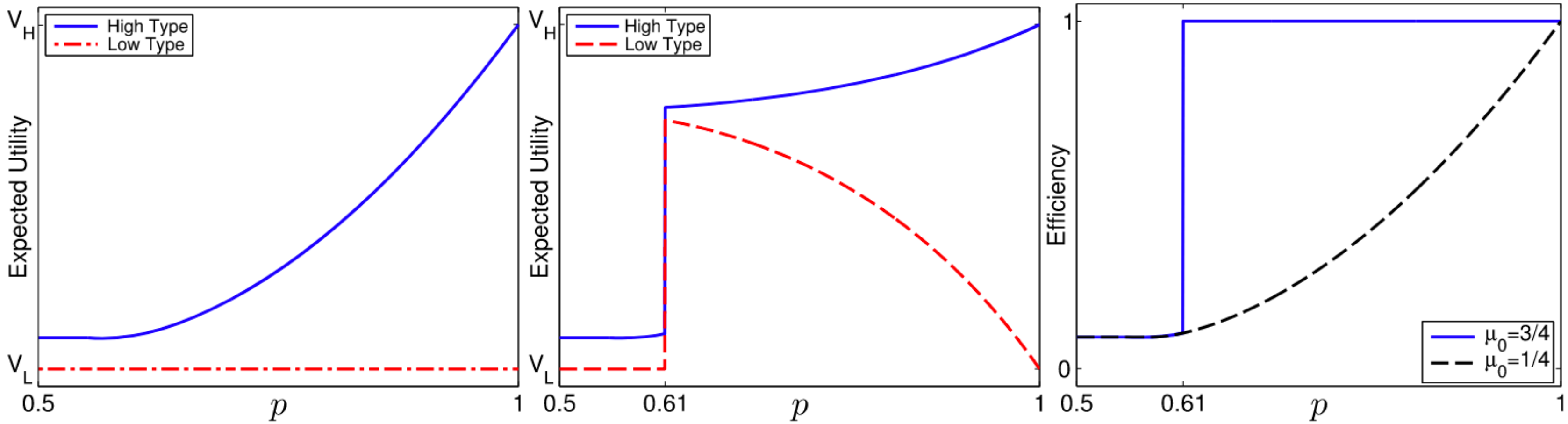


Fig. 5. Comparative statics for x^* and μ^* with symmetric binary tests.

Payoffs.

The high type's welfare increases with the test informativeness, but may increase or decrease with his cost advantage. The low type's welfare decreases with the cost advantage, but may increase or decrease with the test informativeness.



(a) Sender payoff given $\mu_0 = 1/4$ (b) Sender payoff given $\mu_0 = 3/4$ (c) Efficiency

Fig. 6. Sender welfare (a), (b) and efficiency (c) as they depend on the informativeness of the symmetric binary test for two different priors. All panels use $C_L/C_H = 1.1$.

Efficiency.

Efficiency increases with the test informativeness, but in a way that depends on the prior. If the prior is low, efficiency increases continuously with test informativeness and approaches first-best only as the test becomes perfect. If the prior is high, the first-best outcome can be achieved with an imperfect test.

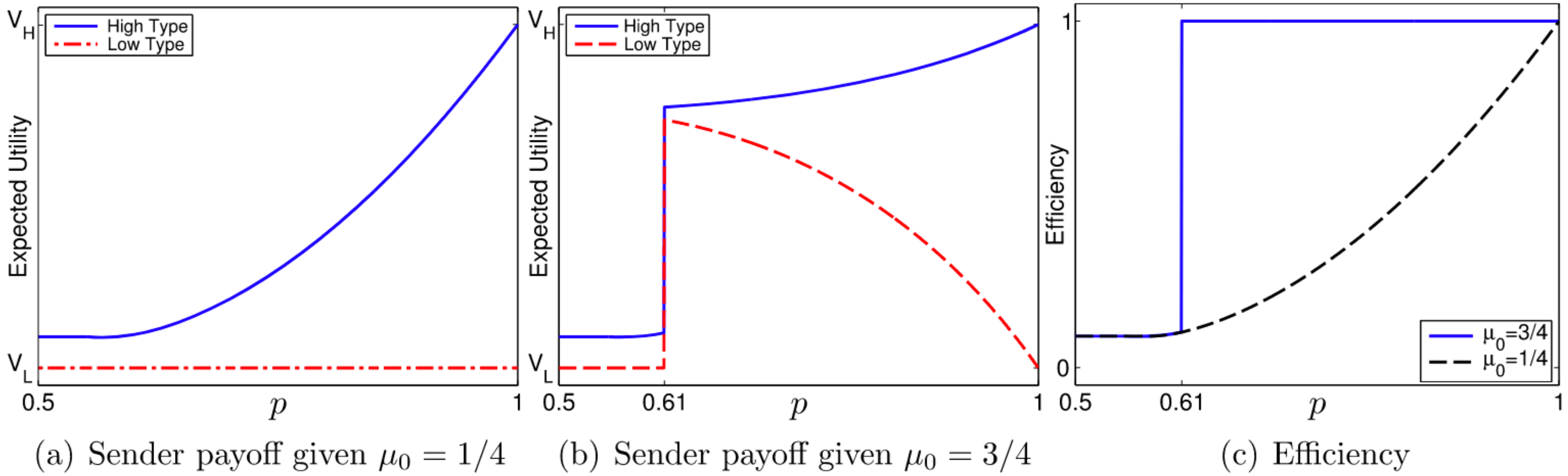


Fig. 6. Sender welfare (a), (b) and efficiency (c) as they depend on the informativeness of the symmetric binary test for two different priors. All panels use $C_L/C_H = 1.1$.

A.1 $\frac{-U_{H,1}}{-U_{H,2}} \leq \frac{-U_{L,1}}{-U_{L,2}}$ for all (x, μ_f).

A.2 $U_t(x, \mu_{t'})$ is strictly quasiconcave in x for all t, t', where μ_t denotes the degenerate belief that places probability one on type t .

A.3 There exists an $\hat{x} \geq 0$ and $d > 0$ such that $U_{t,1}\left(x, \mu_f\right) < -d$ for all t, μ_f and $x \geq \hat{x}$.

A.4 $x_H^* < x^{LCSE}$.

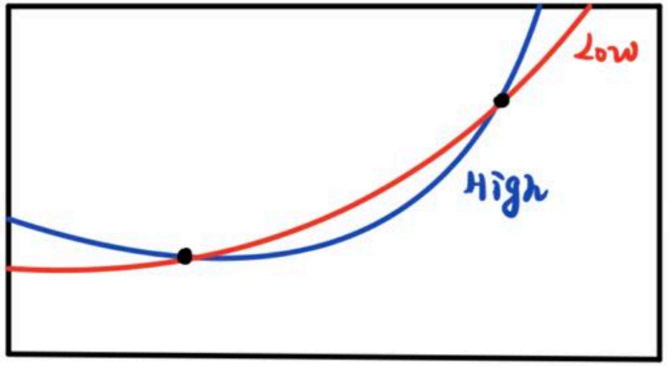
4 Generalized preferences and testing technologies

	Standard signaling with grades	Generalized preferences and testing technologies
Measurement of grade g	$R(g) \equiv \frac{f_L(g)}{f_H(g)}$	$R(g\|x) = \frac{f_L(g\|x)}{f_H(g\|x)}$
RC-informative	$E[R(g)\|L] > \frac{C_L}{C_H}$	$E[R(g\|x)\|L,x] > \frac{U_{L,1}(x,\mu_f)}{U_{L,2}(x,\mu_f)} / \frac{U_{H,1}(x,\mu_f)}{U_{H,2}(x,\mu_f)}\|_{\mu_f=1}$
Expected utility of the type t worker	$u_t(x,\mu)$=$w_t(\mu)$-$C_t x$	$\mu_t(x,\mu) = \int U_t(x,\mu_f(\mathrm{x},\mathrm{g}))f_t(g\|x)dg$
D1 equilibria	If the test is RC-Informative ,D1 equilibria depends on the comparison of μ_0 and μ^*.	If the test is RC-Informative at x^{LCSE}, then for any $\mu_0 \in (0,1)$, all D1 equilibria involve some degree of pooling.
Convergence	If the test is RC-Informative, then as $\mu_0 \to 1$, the D1 equilibria → complete-information.	If the test is RC-Informative at all $x \in [x_H^*, x^{LCSE}]$, then as $\mu_0 \to 1$, the D1 equilibria → complete-information.
BIC Property	Single-crossing property	Double-crossing property (DCP)

Double-crossing property (DCP).

Consider any feasible $\widehat{u_L}, \widehat{u_H}$ such that $b_L(x_0|\widehat{u_L}) = b_H\ (x_0|\widehat{u_H})$ for some $x_0 \in [0, \bar{x}(\widehat{u_L})$.
If $\frac{\partial}{\partial x} b_H(x_0|\widehat{u_H}) \leq \frac{\partial}{\partial x} b_L(x_0|\widehat{u_L})$, then $b_H(x_0|\widehat{u_H}) > b_L(x_0|\widehat{u_L})$ at all $x < x_0$.

- DCP says that:
  - if the high type's BIC is flatter than the low type's at a point of intersection, then the high type's BIC lies everywhere above the low type's at all points to the left.
  - if the low type's indifference curve is flatter at a point of intersection, then it lies everywhere below to the right.
- DCP is implied by MRSP, but not the converse. The property arises naturally on BICs in many signaling models with grades.



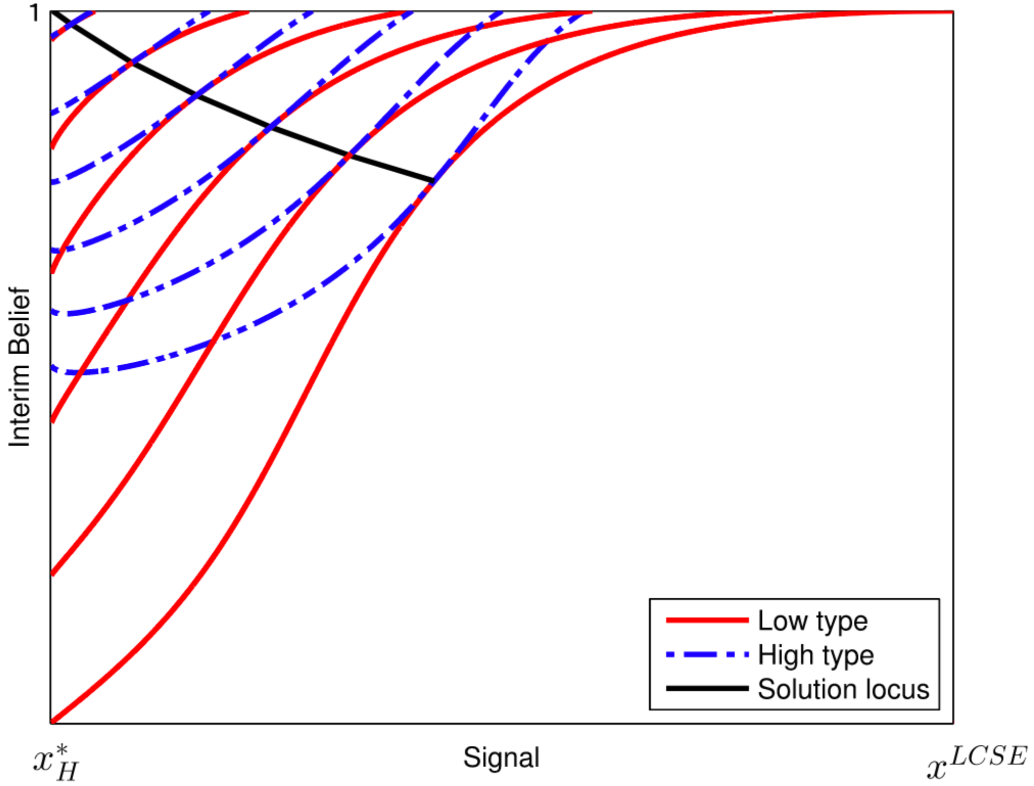
Proposition 4.4

Under DCP, if the test is not RC-Informative at x^{LCSE}, then for all $\mu_0 \in (0,1)$, the LCSE is a D1 equilibrium.

Proposition 4.5

Under DCP, if there exists an $x \in [x_H^*, x^{LCSE}]$ such that the test is not RC-Informative at x, then as $\mu_0 \to 1$, the set of D1 equilibria does not converge to the complete-information outcome by either notion of convergence (type-by-type or in total mass).

- Properties of the solution locus depend on the informativeness of the testing technology.
  - If the test is RC-Informative: the high type maximizes expected utility by relying at least partially on the outcome of the test.
  - If the test is not RC-Informative: the high type maximizes his expected utility by completely separating from the low type using the costly action.



Proposition 4.6

Suppose that DCP holds and $\mu_H(\widehat{u_L})$ is non-decreasing. Then, there exists a unique D1 equilibrium for almost all priors, $\mu_0 \in (0,1)$.

Generically, if $\mu_0 \leq \mu_H(\underline{u_L})$ the D1 equilibrium is:

- $\sigma_H^*\left(x_H\left(\underline{u_L}\right)\right) = 1$, $\sigma_L^*\left(x_H\left(\underline{u_L}\right)\right) = \frac{1-\mu_H(\underline{u_L})}{\mu_H(\underline{u_L})}\frac{\mu_0}{1-\mu_0}$, and $\sigma_L^*(x_L^*) = 1 - \sigma_L^*\left(x_H\left(\underline{u_L}\right)\right)$,
- $\mu\left(x_H\left(\underline{u_L}\right)\right) = \mu_H(\underline{u_L})$, and $\mu(\mathrm{x}) = 0$ if $x \neq x_H\left(\underline{u_L}\right)$.

If $\mu_0 > \mu_H(\underline{u_L})$, the D1 equilibrium is full pooling at $x_H(u_L^*)$, where u_L^* satisfies $\mu_H(u_L^*) = \mu_0$.

5.1 Applications

- 1. Advertising: Kihlstrom and Riordan (1984)

The test in this example represents product reviews, such as those provided by Yelp, CNET, Zagat, Angie’s List, etc.

- 2. Warranties: AsinGal-Or

The test in this example is conducted by third-party reviewers such as Consumer Reports, J.D. Power and Associates, etc.

- 3. Financial structure and inside information: Leland and Pyle

The grade is an analyst’s recommendation, buy or sell, prior to the issuance date, which is the outcome of a symmetric binary.

- 4. Auditors and equity issuance:

The auditor then prepares a statement, which serves as the grade.

Relevance of grades.

The presence of an informative test implies that grades convey meaningful information in equilibrium because the predicted behavior changes from separation to pooling on the level of costly signaling.

Test precision.

Less informative grades increase the amount of resources devoted to inefficient signaling activities.

Reputation matters for signaling.

Signaling behavior depends on the initial market belief about the sender; a sender with a better reputation incurs lower signaling costs.

Signaling without single-crossing.

In the presence of grades, the sender can (imperfectly) signal his type through money burning activities.

5.3 Extensions

- In many relevant applications, the sender may have channels through which he has influence over the outcome of the test.
  - affect the outcome of the test directly by undertaking some costly action;
    - a student chooses how much effort to exert studying prior to an exam
  - indirectly by improving the distribution of the quality being tested
    - a firm chooses how much to invest in product quality

5.3.1 Indirect influence: Ex-ante investment

- Suppose that, before learning his type, the sender chooses how much to "invest" in his type: the more he invests, the more likely he is to become the high type.
- There are two important benchmarks: the investment level that arises
  - if the type is perfectly revealed to receivers
  - if there is no grade
- An important consideration is whether the sender's exante investment is observable.
  - When investment is unobservable, the sender always invests less than in the complete-information/first-best benchmark. And investment can be higher or lower than the no-grade benchmark.
  - Observability strengthens the incentives to invest.

5.3.2 Direct influence: Ex-interim hidden effort

- After the sender chooses x, but before the grade is realized, the sender chooses a level of costly, unobservable effort toward improving his grade on the test. The game can be thought of as a combination of our model and a noisy signaling model.
- This extension serves largely as a check that the model is robust to a realistic feature of some environments.
- The main substantive implication it generates is that the observable action and hidden effort are, to a degree, substitutes.

6 Conclusion

- Grades are a prevalent force in many incomplete-information environments and often convey meaningful information. There is a subtle interaction between the information conveyed by costly observable actions and information conveyed by grades.
  - If the test is sufficiently informative, the high type relies on its ability to convey information, and the stable equilibrium involves some degree of pooling.
  - We have used D1 to refine the set of PBE within our model.
- In many examples, such as education, the costly signaling action involves waiting to trade. Static signaling models, such as the one presented here, ignore the dynamic aspects of such an environment.
  - They dynamic model relaxes the assumption of a seller's ability to commit to delay trade until a fixed date x. From this standpoint, one could investigate how the timing of information revelation interacts with commitment power to impact trading patterns and welfare.

Thank You !