A sufficient condition for all-or-nothing information supply in price discrimination*

Marco LiCalzi

Department of Applied Mathematics – University of Venice
30123 Venezia, Italy
licalzi@unive.it
http://helios.unive.it/licalzi

Abstract. This paper provides a sufficient condition under which the optimal policy of a monopolistic seller who is considering the tradeoff between price discrimination and information disclosure is at one of two extremes: either buyers are given access to all the available information, or the seller makes no disclosure at all.

Keywords. Comparative statics, supermodular functions, lattice programming.

M.S.C. classification: 91B44, 90C56, 06B99.
J.E.L. classification: C69, D82.

1 Introduction

Increasing revenue by means of price discrimination is a common practice among monopolistic sellers. Lewis and Sappington (1994) explores the seller's tradeoff between information disclosure and price discrimination. On one hand, disclosing information to prospective buyers enables the seller to segment the market and charge higher prices to high-value buyers. On the other hand, improved information may help buyers to earn informational rents which damages seller's profits. A few examples in Lewis and Sappington (1994) show that in a variety of settings this tradeoff is optimally resolved with an all-or-nothing information disclosure policy: either buyers should be supplied with the best available knowledge, or the seller should not disclose any information at all.

A counterexample in Lewis and Sappington (1994) shows that this conclusion does not hold in general, but that paper fails to provide any general sufficient condition for its validity. The modest purpose of this paper is to provide one such condition, which is based on four assumptions. This sufficient condition is derived using the techniques of monotone comparative statics. Very recently, Johnson and Myatt (2004) have used a different approach to provide a more

* We thank Marco Ottaviani for raising our interest in this problem. Financial support from MIUR is acknowledged.

general alternative condition. However, in spite of its formal elegance, the abstract nature of their approach takes priority on its intuitive content. Therefore, we provide here a second more stringent sufficient condition which replaces two of our original four technical assumptions with hypotheses that enjoy a simple and natural economic interpretation.

2 The problem

There is a (male) monopolistic seller who wishes to sell one indivisible object to one[1] (female) buyer by means of a take-it-or-leave-it offer at price p. Both agents are risk-neutral. The buyer does not know the value $V \geq 0$ of the object, but she can observe a signal S of type θ. Intuitively, we think of θ as a measure of the "quality" of the signal S: the higher θ, the more accurate is the information that S carries about V. The seller can (credibly and costlessly) control the quality θ of the signal observed by the buyer before she takes her decision whether to purchase or not. We assume that θ can vary within a compact interval $\Theta = [\theta_L, \theta_H]$ with $\theta_L \geq 0$ and that V and S are independent whenever $\theta = 0$. Thus, a signal S of quality $\theta = 0$ carries no information about V.

In this situation, the buyer's optimal choice is the following. After receiving a signal s of quality θ, she buys the object whenever its expected value is not lower than the price p demanded by the seller. Let $b(s, \theta) = E(V|s, \theta)$ be the expected value of V conditional on receiving a signal $S = s$ of quality θ. Then the buyer purchases whenever $b(s, \theta) \geq p$.

Now, consider the seller's problem. Suppose that the seller has an opportunity cost c for the object. Then his profit from selling the object is

$$\Pi(p, \theta) = (p - c) \cdot P[b(s, \theta) \geq p].$$

Knowing the buyer's optimal behavior, the seller wishes to choose a pair (p^*, θ^*) which maximizes his expected profit. Clearly, the optimal price $p^*(\theta)$ depends in general on θ but for simplicity we usually omit the argument.

This paper provides a sufficient condition such that either θ_L or θ_H are always part of the optimal seller's choice. That is, it is part of the optimal policy for the seller to use an all-or-nothing policy which makes the quality of the signal extreme.

The problem of providing sufficient conditions for an all-or-nothing information disclosure policy can be attacked in several ways. One is by starting out with assumptions on the joint distribution $H(v, s)$ for the buyer's value $V \geq 0$ and the signal S. Another one is to begin with assumptions on the prior distribution $F(v)$ on V and the conditional distribution $G(s|v)$ for the signal (given the value V). The third one is to work with the family of conditional distributions $F(v|s)$, and this is the route that we follow here. The three approaches are strictly related and one may easily recast the assumptions differently, although the choice of the

[1] An equivalent formulation assumes a unit mass of consumers with valuation θ. See Johnson and Myatt (2004).

approach impacts on the naturalness of the hypotheses. Another possible line of attack is to find sufficient conditions for the stronger result that $\Pi(p^*(\theta), \theta)$ is quasiconvex in θ. Appendix A.2 provides an explicit example in this respect.

3 The main result

Denote by $H(v, s; \theta)$ the bivariate joint distribution for the random vector $(V, S)_\theta$ when the quality of the signal is θ. Recall that $b(s, \theta)$ is the expectation for the buyer's value conditional on her receiving a signal $S = s$ of quality θ. Our set of sufficient conditions requires four assumptions.

(A1) For any θ, the distribution $H(v, s; \theta)$ has the same marginals, denoted respectively by $F(v)$ and $G(s)$.

This is a structural assumption that neither requires V or S to have connected support, nor H (or F or G) to be absolutely continuous. The only restriction imposed is the very mild property that θ does not affect the prior distributions for V and S. To avoid trivialities, in the following we assume also that both the support of V and the support of S contain at least two distinct points.

The second assumption concerns the relationship between V and S (regardless of θ). Given the random vector (V, S), denote by $F(v|s)$ the distribution of V conditional on $S = s$. We say that V is stochastically increasing in S if

$$F(v|s_1) \geq F(v|s_2) \qquad \text{for all } s_1 \leq s_2 \text{ and } v; \tag{1}$$

that is, if $s_1 \leq s_2$ implies that $F(v|s_1)$ is stochastically dominated by $F(v|s_2)$. This assumption was introduced first in Lehmann (1966) under the name of "positive regression dependence," but it is nowadays known also as "stochastic monotonicity" or "conditional stochastic dominance".

(*) For any θ, the value V is stochastically increasing in the signal S.

It is well known that affiliation of V and S implies (*). The opposite does not hold even if we assume that the joint density of (V, S) exists; see Lewis and Sappington (1988, p. 426) for a counterexample. It is also known that (A2) implies a few weaker notions of positive dependence between V and S, and in particular positive quadrant dependence: $H(v, s) \geq F(v) \cdot G(s)$ for all (v, s); see Nelsen (1999, p. 163). For simplicity, we assume as much differentiability as convenient and thus the equivalent formulation of (*) becomes the following. (We also leave it understood that inequalities such as $F_s > 0$ are supposed to hold over the entire domain of the function.)

(A2) The conditional distributions $F(v|s; \theta)$ are twice continuously differentiable functions with $F_s > 0$.

The third assumption requires that increasing θ reinforces the strength with which the main clause of (*) is satisfied. We make a short detour to motivate this statement. Given the conditional distribution $F(v|s)$, denote by $F^{-1}(\cdot|s)$ its right-continuous inverse (with respect to v). The property of stochastic monotonicity defined in (1) is equivalent to

$$F^{-1}(y|s_1) \leq F^{-1}(y|s_2) \qquad \text{for all } s_1 \leq s_2 \text{ and all } y \text{ in } [0,1].$$

Based on this, Yanagimoto and Okamoto (1969) defined a random vector $(V,S)_{\theta_2}$ to be "more stochastically increasing" than another random vector $(V,S)_{\theta_1}$ whenever, for any $s_1 \leq s_2$,

$$F^{-1}(y|s_1,\theta_1) \leq F^{-1}(z|s_2,\theta_1) \Rightarrow F^{-1}(y|s_1,\theta_2) \leq F^{-1}(z|s_2,\theta_2) \tag{2}$$

for all y, z in $[0,1]$. Yanagimoto and Okamoto (1969) shows that this definition is equivalent to requiring that, for any $s_1 \leq s_2$,

$$F(v|s_1,\theta_1) \geq F(w|s_1,\theta_2) \Rightarrow F(v|s_2,\theta_1) \geq F(w|s_2,\theta_2) \tag{3}$$

for all v, w in $[0,1]$. Using a subscript[2] to denote the dependence from θ, this can be equivalently written as

$$F_{\theta_2}^{-1}\left[F_{\theta_1}(v|s_1)\Big|s_1\right] \leq F_{\theta_2}^{-1}\left[F_{\theta_1}(v|s_1)\Big|s_2\right] \qquad \text{for all } s_1 \leq s_2 \text{ and } v. \tag{4}$$

Our third assumption requires that (3) holds for any $\theta_1 \leq \theta_2$; that is, raising θ must increase the stochastic monotonicity of $(V,S)_\theta$. We note that exchanging v and s in (4) yields the definition of "increasing accuracy" in the signal proposed in Lehmann (1988) and adopted in Persico (2000). Based on this, we dub (3) as a property of increasing "accuracy" in the value V or, for short, V-accuracy.

(**) The family $(V,S)_\theta$ is increasingly V-accurate in θ.

Both increasing V-accuracy and the symmetric notion of increasing S-accuracy imply that the family $(V,S)_\theta$ is increasingly concordant in θ; that is, $H(v,s;\theta_1) \leq H(v,s;\theta_2)$ as $\theta_1 \leq \theta_2$. See Thm. 2.12 in Joe (1997, p. 43).

Assuming differentiability and using the obvious convention for the case $F_v = 0$, an equivalent formulation for (**) is the following.

(A3) The conditional distributions $F(v|s;\theta)$ satisfy the Spence–Mirrlees condition:

$$\frac{F_\theta}{F_v} \text{ is decreasing in } s.$$

To see why, observe that Theorem 2.14 in Joe (1997, p. 44) establishes that a sufficient condition for increasing V-accuracy is

$$\frac{\partial^2 F}{\partial v \partial s}\frac{\partial F}{\partial \theta} - \frac{\partial^2 F}{\partial \theta \partial s}\frac{\partial F}{\partial v} \geq 0, \tag{5}$$

[2] Anywhere else in this paper, the notation F_θ stands for the partial derivative of F with respect to θ.

and that this holds if and only if F_θ/F_v is decreasing in s.

We remark that the Spence–Mirrlees condition provides a particularly simple condition to check. Moreover, Appendix A.1 shows that, under a very mild regularity condition, it is in fact necessary and sufficient for (**).

The fourth and last assumption requires that there exists a signal which is neutral with respect to quality.

(A4) There exists a signal $\hat{s}$ such that $b(\hat{s}, \theta)$ is constant for all θ.

Our main result is that, jointly taken, these four assumptions imply the optimality of an extreme policy for information disclosure. Before giving the main result, we prove another (slightly more general) sufficient condition. The assumptions have a technical flavor and are not meant to be intuitive. This proposition, on the other hand, provides conditions which are easier to check in practice and makes the connection with the standard techniques in monotone comparative statics more transparent.

Proposition 1. *Assume (A1) and (A4) and suppose that the following holds:*

(B2) *The function $b(s, \theta) = E(V|s, \theta)$ is continuously differentiable and $b_s > 0$.*

(B3) *The function $b(s, \theta)$ is supermodular.*

Then it is part of the optimal policy for the seller to make the quality of the signal extreme.

Proof. It suffices to show that if (p^*, θ^*) is an optimal choice for the seller but $(\theta^* - \theta_L)(\theta^* - \theta_H) \neq 0$ then

$$\Pi(p^*, \theta^*) \leq \max \left\{ \Pi(p^*, \theta_L), \Pi(p^*, \theta_H) \right\}. \tag{6}$$

Let $\hat{s}(\theta)$ be the implicit function defined by $b(s, \theta) = p^*$. By (B2), the function $\hat{s}(\theta)$ is defined for all θ and, moreover, $P[b(s, \theta) \geq p^*] = P[S \geq \hat{s}(\theta)]$.

Recall that $\Pi(p, \theta) = (p - c) \cdot P[b(s, \theta) \geq p]$. Since $(p^* - c) \geq 0$, $\Pi(p^*, \theta)$ is an increasing function of $P[b(s, \theta) \geq p^*] = P[S \geq \hat{s}(\theta)]$. By (A1), $P[S \geq \hat{s}(\theta)]$ is a decreasing function of $\hat{s}(\theta)$. Therefore, (6) follows if we show that θ^* cannot be a strict local minimizer for $\hat{s}(\theta)$ when $\theta_L < \theta^* < \theta_H$.

By contradiction, suppose that θ^* is a strict local minimizer for $\hat{s}(\theta)$ and that $\theta_L < \theta^* < \theta_H$. Then there exist θ_1 and θ_2 such that $\theta_L \leq \theta_1 < \theta^* < \theta_2 \leq \theta_H$ and $\hat{s}(\theta^*) < \min \{\hat{s}(\theta_1), \hat{s}(\theta_2)\}$. Let $s^* = \hat{s}(\theta^*)$ and $s_i = \hat{s}(\theta_i)$ for $i = 1, 2$. Without loss of generality, suppose $s_1 \leq s_2$. Using (B2) twice, $b(s_1, \theta^*) > b(s^*, \theta^*) = b(s_1, \theta_1) = b(s_2, \theta_2) \geq b(s_1, \theta_2)$. Therefore, $b(s_1, \theta^*) - b(s_1, \theta_1) > 0 \geq b(s_1, \theta_2) - b(s_1, \theta^*)$, which — together with (A4) — contradicts the supermodularity of b.

We are now ready to state the main result of this section.

Proposition 2. *Assume (A1), (A2), (A3), and (A4). Then it is part of the optimal policy for the seller to make the quality of the signal extreme.*

Proof. By Proposition 1, it suffices to show that (A2) and (A3) imply (B2) and (B3). Clearly, (B2) follows immediately from (A2). As for (B3), let $W \subset \mathbb{R}_+$ denote the support of V. Observe that

$$\begin{aligned} b_\theta(s,\theta) &= \frac{\partial}{\partial\theta}\int_W [1 - F(v|s,\theta)]\;\mathrm{d}v = -\int_W [F_\theta(v|s,\theta)]\;\mathrm{d}v \\ &= -\int_W \left[\frac{F_\theta}{F_v}\right] F_v(v|s,\theta)\,\mathrm{d}v. \end{aligned}$$

The Spence–Mirrlees condition implies that the last integral is a convex combination with respect to $F_v(v|s,\theta)$ of functions which are decreasing in s. Hence, b_θ is increasing in s or, equivalently, $b(s,\theta)$ is supermodular.

4 Examples

This section illustrates the power of Proposition 1 with a few examples. We believe that they have independent interest for the fields of economics of information because they show how copulas can be conveniently used to produce parametric examples in problems with asymmetric information. Since lack of space prevents us from discussing copulas here, the interested reader is referred to Nelsen (1999).

Recall that $H(v,s;\theta)$ denotes the bivariate joint distribution for the random vector $(V,S)_\theta$ when the quality of the signal is θ. Moreover, (A1) postulates that, for any θ, $H(v,s;\theta)$ has identical marginals $F(v)$ and $G(s)$. Applying the theory of copulas, this implies that there exists a copula $C^\theta(x,y)$ on $[0,1]\times[0,1]$ such that $H(v,s;\theta) = C^\theta[F(v),G(s)]$.

Stochastic independence is associated with the product copula $C^0(x,y) = xy$. Given a copula $C(x,y)$, we define its "excess" with respect to the product copula by the function $\mathcal{E}(x,y) = C(x,y) - C^0(x,y)$. We assume for simplicity that $\mathcal{E}$ is differentiable. (A mild generalization is possible under the assumption that $\mathcal{E}$ be absolutely continuous.) By the elementary properties of copulas, any excess function $\mathcal{E}(x,y)$ has $\mathcal{E}(0,y) = \mathcal{E}(x,0) = \mathcal{E}(1,y) = \mathcal{E}(x,1) = 0$ and satisfies $\mathcal{E}_{xy}(x,y) \geq -1$ for all (x,y).

We say that a copula $C(x,y)$ is *informative* when C is concave in y for all x. As it turns out, (A2) implies that C is informative and this implies that it satisfies positive quadrant dependence, which is equivalent to $\mathcal{E}(x,y) \geq 0$ for all x,y.

4.1 Separable excess

In this section, we assume that the excess function $\mathcal{E}(x,y) = C(x,y) - xy$ is multiplicatively separable. That is, there exist two twice continuously differentiable functions Γ and Φ such that $\mathcal{E}(x,y) = \Gamma(x)\Phi(y)$. Denote by γ and φ their respective derivatives. In order to guarantee that $\mathcal{E}(x,y)$ be an excess function, we assume

(E1) Γ and Φ are two positive functions on $[0,1]$.
(E2) $\Gamma(0)=\Gamma(1)=\Phi(0)=\Phi(1)=0$.
(E3) $|\gamma(v)|\leq 1$ and $|\varphi(s)|\leq 1$.

The resulting copula is of the form

$$C(x,y)=xy+\Gamma(x)\Phi(y).$$

Note that (A1–A3) imply $0\leq\Gamma(x)\leq\min\{x,1-x\}$. Finally, to ensure that the copula be informative and to avoid trivialities, we assume

(E4) $\Phi(s)$ is concave, with $\phi'<0$.

Proposition 3. *Let $\mathcal{C}$ be a family of copulas with separable excess that satisfy (E1), (E2), (E3), and (E4). For any copula C in $\mathcal{C}$, let*

$$\theta=\int_0^{+\infty}\Gamma[F(v)]\,dv.$$

Then an extreme quality of the signal is optimal.

Proof. We apply Proposition 1. Clearly, (A1) holds by assumption. To check the other assumptions, let us compute $b(s,\theta)$. Given that the joint distribution of (V,S) is $H(v,s)=F(v)\cdot G(s)+\Gamma[F(v)]\cdot\Phi[G(s)]$, the conditional distributions are of the form $F(v|s,\theta)=F(v)+\Gamma[F(v)]\cdot\varphi[G(s)]$. Hence,

$$b(s,\theta)=E(V|s,\theta)=\int_0^{+\infty}[1-F(v|s,\theta)]\;\mathrm{d}v=E(V)-\theta\varphi[G(s)].$$

Since $b_s=-\theta\varphi'[G(s)]g(s)$, (E4) implies (B2). Similarly, $b_{s\theta}=-\varphi'[G(s)]g(s)>0$ implies (B3). Finally, (E2) and (E4) imply that ϕ changes sign from positive to negative as s increases. By continuity, there exists s' such that $\phi(s')=0$. This implies (A4) and concludes the proof.

4.2 Convex combinations

For our second example, we consider the family of copulas defined by

$$C^{\theta}(x,y)=xy+\theta\mathcal{E}(x,y), \tag{7}$$

where $\mathcal{E}(x,y)$ is an excess function and θ is in $[0,1]$. The major point of interest for this family is that, if we let $C^1(x,y)=\mathcal{E}(x,y)-xy$ be an informative copula, the copulas in (7) can be written as

$$C^{\theta}(x,y)=\theta C^0(x,y)+(1-\theta)C^1(x,y). \tag{8}$$

Thus, as θ increases, we move away from the (noninformative) product copula $C^0(x,y)=xy$ to the informative $C^1(x,y)$.

In the following, we assume that the excess function $\mathcal{E}(x,y)$ is differentiable and concave in y for all x, with $\mathcal{E}_{yy}<0$. These assumptions are made mostly for convenience, except for the concavity assumption which ensures that C^1 is informative.

Proposition 4. *For the family of convex combinations of the product copula and of an informative copula, an extreme quality of the signal is optimal.*

Proof. Given the copula, the conditional distributions are of the form $F(v|s,\theta) = F(v) + \theta\mathcal{E}_s[F(v), G(s)]$. Define

$$\varphi(s) = \int_0^{+\infty} [1 - F(v|s,\theta)]\ \mathrm{d}v.$$

Note that $\varphi(s)$ is decreasing with $\phi' < 0$. Let

$$b(s,\theta) = E(V|s,\theta) = \int_0^{+\infty} [1 - F(v|s,\theta)]\ \mathrm{d}v = E(V) - \theta\varphi[G(s)].$$

From here on, the proof is identical to the one offered for Proposition 3.

4.3 The normal example

This example has a different structure. We suppose that V is normally distributed with mean μ and precision τ_V; for short, $V \sim N(\mu, \tau_V)$. The signal S is given by $S = V + \varepsilon$, where the noise term ε is normally distributed with zero mean and precision τ_ε; that is, $\varepsilon \sim N(0, \tau_\varepsilon)$.

Proposition 5. *For the normal example, an extreme quality of the signal is optimal.*

Proof. It is well-known that

$$E(V|s,\tau_\varepsilon) = \mu + \frac{\tau_\varepsilon(s-\mu)}{\tau_\varepsilon + \tau_V}.$$

Let $\theta = \tau_\varepsilon/(\tau_\varepsilon + \tau_V)$ and $\varphi(s) = (s - \mu)$. Then we obtain

$$b(s,\theta) := E(V|s,\theta) = E(V) - \theta\varphi(s).$$

From here on, the proof is identical to the one offered for Proposition 3.

4.4 Plackett's family

For this example, we consider the Plackett's family of copulas defined by

$$C^\theta(x,y) = \frac{[1 + \theta(s+v)] - \sqrt{[1+\theta(s+v)]^2 - 4sv\theta(\theta+1)}}{2\theta}, \tag{9}$$

for $\theta > -1$. In Nelsen (1999, p. 160), we find

$$C^\theta_{yy}(x,y) = \frac{-2\theta(\theta+1)x(1-x)}{\left([1+\theta(x+y)]^2 - 4xy\theta(\theta+1)\right)^{3/2}},$$

which is negative for $\theta \geq 0$. Hence, this family exhibits informativeness only for $\theta \geq 0$ and thus we assume that θ varies in (a compact interval) of $\mathbb{R}_+$. Note that $C^0(x, y)$ is the product copula.

We also assume that V is uniformly distributed on $[0, 1]$ so that $F(v) = v$. Carrying out the (long) integration, we obtain

$$b(s, \theta) := E(V|s, \theta) = \left[\frac{(\theta+1)\ln(\theta+1) - \theta}{\theta^2}\right] + \left[\frac{\theta^2 + 2\theta - 2(\theta+1)\ln(\theta+1)}{\theta^2}\right] s.$$

Let

$$\alpha(\theta) = \frac{(\theta+1)\ln(\theta+1) - \theta}{\theta^2} \qquad \text{e} \qquad \beta(\theta) = \frac{\theta^2 + 2\theta - 2(\theta+1)\ln(\theta+1)}{\theta^2}$$

and rewrite

$$b(s, \theta) = \alpha + \beta s,$$

omitting the argument θ for convenience. The function $b(s, \theta)$, is a linear function of s for all θ.

Note that both $\alpha(\theta)$ and $\beta(\theta)$ are functions defined for $\theta > 0$ and can be extended to $\theta = 0$ by continuity. The function $\alpha(\theta)$ is decreasing, convex and onto $(1/2, 0)$; while $\beta(\theta)$ is increasing, concave and onto $(0, 1)$. As θ increases, $b(s, t)$ has a lower intercept but a greater slope: in the limit, $b(s, 0) = 1/2$ and $b(s, +\infty) = s$.

Proposition 6. *For the Plackett's family with V uniformly distributed on $[0, 1]$, an extreme quality of the signal is optimal.*

Proof. Once again, it suffices to apply Proposition 1. Clearly, (A1) holds by assumption. Since $b_s = \beta(\theta) > 0$, (B2) holds. Similarly, $b_{s\theta} = \beta'(\theta) > 0$ implies (B3). Finally, for $s^\circ = 1/2$, $b(s^\circ, \theta) = E(V)$ for all θ which implies (A4) and concludes the proof.

Under the additional assumption that S is also uniformly distributed, Appendix A.2 provides additional computations for the family in this example. These include a derivation of the optimal pair (p^*, θ^*) for any cost c and prove the stronger result that $\Pi(p^*(\theta), \theta)$ is quasiconvex in θ.

A Appendices

A.1 The Spence–Mirrlees condition is equivalent to (**)

Milgrom and Shannon (1994) define a function $F(v|s, \theta)$ to be *completely regular* if it is continuously differentiable and its level sets $\{(v, \theta) : F(v|s, \theta) = k\}$ are path-connected. Complete regularity is implied by $F_v > 0$ and $F_\theta \neq 0$ for all v in the interior of the support of $V|S = s$ and for all θ in the interior of Θ.

Assumption (**) states that

$$F(v|s_1, \theta_1) \geq F(w|s_1, \theta_2) \Rightarrow F(v|s_2, \theta_1) \geq F(w|s_2, \theta_2) \tag{10}$$

for all v, w in $[0,1]$ whenever $\theta_2 \geq \theta_1$. Endow the domain of (v,θ) in $\mathbb{R}^2$ with a lexicographic order $\succeq$ such that $(v_2,\theta_2) \leq (v_1,\theta_1)$ if either $\theta_2 > \theta_1$ or $\theta_2 = \theta_1$ and $v_2 \geq v_1$. Then (10) holds if and only if the function $U(v,\theta,s) = -F(v|s,\theta)$ satisfies the single crossing property in (v,θ): for $(v_2,\theta_2) \succeq (v_1,\theta_1)$ and $s_2 \geq s_1$, $U(v_2,\theta_2,s_1) > U(v_1,\theta_1,s_1)$ implies $U(v_2,\theta_2,s_2) > U(v_1,\theta_1,s_2)$ and $U(v_2,\theta_2,s_1) \geq U(v_1,\theta_1,s_1)$ implies $U(v_2,\theta_2,s_2) \geq U(v_1,\theta_1,s_2)$.

The following proposition is a straightforward application of Theorem 3 in Milgrom and Shannon (1994, p. 161).

Proposition 7. *Suppose that $F(v|s,\theta)$ is completely regular. Then (**) holds if and only if (A3) holds.*

Note that the conclusion holds even if we drop the assumption $F_s > 0$ from (A2).

A.2 Plackett's copula

Let $H(v,s;\theta)$ be the joint distribution of the buyer's value V and of the signal S (when its quality is θ). Given the corresponding marginal c.d.f.'s $F(v)$ and $G(s)$, there exists a copula $C^\theta(x,y)$ such that $H(v,s;\theta) = C^\theta[F(v),G(s)]$. We suppose that this copula belongs to the Plackett's family and that V and S are both uniformly distributed, so that $F(v) = v$ and $G(s) = s$.

The Plackett family of copulas is defined for $t > 0$. (Setting $\theta = t-1$ restores the conventions used in the main text.) For $t \neq 1$ we have:

$$C(v,s|t) = \frac{[1+(t-1)(s+v)] - \sqrt{[1+(t-1)(s+v)]^2 - 4svt(t-1)}}{2(t-1)}$$

and $C(v,s|t) = vs$ for $t = 1$. As t increases, dependence increases. For $t \geq 1$, we have positive dependence. This is a comprehensive family; that is, it includes the copula corresponding to the minimal Frechet distribution ($t \to 0$), the product copula corresponding to the case of stochastic independence ($t = 1$) and copula associated to the maximal Frechet distribution ($t \to +\infty$). See Nelsen (1999, pp. 80–81).

As noted above, Nelsen (1999, p. 160) gives

$$C_{ss}(v,s|t) = \frac{-2t(t-1)v(1-v)}{\left([1+(t-1)(v+s)]^2 - 4vst(t-1)\right)^{3/2}}$$

which is negative for $t \geq 1$. Hence, $C(v,s|t)$ is concave for $t \geq 1$ and the Plackett family satisfies informativeness for $t \geq 1$. From Nelsen (1999, p. 176), we note

$$C_s(v,s|t) = \frac{tv + (1-t)C(v,s|t)}{1+(t-1)[v+s-2C(v,s|t)]}.$$

The conditional density is

$$C_{vs}(v|s,t) = \frac{t[v(1-s)(t-1) + s(1-v)(t-1) + 1]}{[(v-s)^2t^2 + 2(v-v^2+s-s^2)t + (v+s-1)^2]^{3/2}}$$

Integrating with respect to v, we find

$$b(s,t) = E(V|s,t) = \int_0^1 vC_{vs}(v|s,t)\,\mathrm{d}v = \frac{1+t\ln t - t}{(t-1)^2} + \frac{t^2-2t\ln t-1}{(t-1)^2}s.$$

Let

$$A(t) = \frac{1+t\ln t-t}{(t-1)^2} \qquad \text{e} \qquad B(t) = \frac{t^2-2t\ln t-1}{(t-1)^2}$$

and rewrite $b(s,t) = A + Bs$, omitting the argument t for convenience. The function $b(s,t)$, which is the expected value of V given s, is a linear function of s, for all t.

Both functions $A(t)$ and $B(t)$ are defined for $t > 1$ and can be extended to $t = 1$ by continuity. The function $A(t)$ is decreasing, convex and onto $(1/2, 0)$; while $B(t)$ is increasing, concave and onto $(0, 1)$. As t increases, $b(s,t)$ has a lower intercept but a greater slope: in the limit, $b(s, 0) = 1/2$ and $b(s, +\infty) = s$.

Recall that $P(B \text{ buys}) = P(b(s,t) \geq p)$. Using the linearity of $b(s,t)$ in s and the uniform distribution of s,

$$P(b(s,t) \geq p) = P(A + Bs \geq p) = P\left(s \geq \frac{p-A}{B}\right) = 1 - \frac{p-A}{B},$$

provided that the probability remains in $[0, 1]$. (We omit the argument of A and B for simplicity.)

For a given t, the seller's problem becomes

$$\max_p\,(p-c)P(B \text{ buys}) = \max_p\,(p-c)\left[1 - \frac{p-A}{B}\right] \tag{11}$$

This function is concave in p and thus the first order condition suffices to select as the (candidate) global maximizer the price

$$p^\circ(t) = \frac{A+B+c}{2}.$$

However, we also need to impose some conditions to make sure that production and sale at p° is an optimal choice. Recall that the seller may choose to not to sell, in which case he does incur the cost c and makes a zero profit. Hence, the profitability constraint is that $p^\circ(t) \geq c$. Moreover, we need to control for the probability to remain in $[0, 1]$. The probability is not negative if $p \leq A(t) + B(t)$ and it is not over 1 if $p \geq A(t)$. This gives us three constraints to consider.

If

$$\max\{c, A(t)\} \leq p^\circ(t) \leq A(t) + B(t) = \frac{t^2 - t\ln t - t}{(t-1)^2} \tag{12}$$

all constraints are satisfied and p° is indeed the global maximizer. If

$$p^\circ(t) > A(t) + B(t) = \frac{t^2 - t\ln t - t}{(t-1)^2}, \tag{13}$$

the probability of a sale is 0 and therefore the optimal choice is no sale. If $p^\circ(t) < A(t)$, the probability of a sale is 1, hence it is preferable to choose the higher price $p^* = A(t)$, which is more profitable but does not change the probability of sale. If furthermore $A(t) \geq c$, this price is really profitable and thus it is optimal; otherwise, no sale is the optimal choice.

Summarizing, the optimal price p^* (taking into account the constraints) is

$$p^*(t) = \begin{cases} \dfrac{A(t)+B(t)+c}{2} & \text{if } \max\{c, A(t)\} \leq p^\circ(t) \leq A(t)+B(t) \\ A(t) & \text{if } c \leq p^\circ < A(t) \\ +\infty \text{ (no sale)} & \text{if } p^\circ < \min\{c, A(t)\} \text{ or } p^\circ > A(t)+B(t) \end{cases}$$

Expliciting the dependence on t, this gives

$$p^*(t) = \begin{cases} \dfrac{1}{2}\dfrac{t^2 - t\ln t - t}{(t-1)^2} + \dfrac{1}{2}c & \text{if } \max\{c, \dfrac{1+t\ln t - t}{(t-1)^2}\} \leq p^\circ(t) \leq \dfrac{t^2 - t\ln t - t}{(t-1)^2} \\ \dfrac{1+t\ln t - t}{(t-1)^2} & \text{if } c \leq p^\circ < \dfrac{1+t\ln t - t}{(t-1)^2} \\ +\infty \text{ (no sale)} & \text{otherwise} \end{cases}$$

Before computing the profit $\Pi(p^*(t), t)$, let us dub the three possibilities as Case 1, 2, and 3, respectively, and study for which values of c they come in place.

• Case 1 ($\max\{c, A(t)\} \leq p^\circ(t) \leq A(t) + B(t)$). Let us examine this constraint more closely. First, note that two subcases are possible: $A(t) \leq c$ and $A(t) > c$. If $A(t) \leq c$, substitute $p^\circ = (A + B + c)/2$ and the major constraint reads $A \leq c \leq A + B$. If instead $A(t) > c$, the major constraint reads $A - B \leq c < A$. Hence, Case 1 occurs when $A - B \leq c \leq A + B$, or for intermediate values of c. When this restriction fails, we fall back to Case 2 or 3.

The function (defined only for $t > 1$ and extended to $t = 1$ by continuity)

$$A(t) - B(t) = \frac{2 + 3t\ln t - t - t^2}{(t-1)^2}$$

is convex, decreasing and onto $(1/2, -1)$; moreover, it is zero for $t \approx 2.817$. The other function (defined only for $t > 1$ and extended to $t = 1$ by continuity)

$$A(t) + B(t) = \frac{t^2 - t\ln t - t}{(t-1)^2}$$

is concave, increasing and onto $(1/2, 1)$. Together, these two functions define the region of values that c can take while making Case 1 true. See Figure 1. Case 1 is the grey region in the middle. This region is a singleton for $t = 1$ and enlarges itself as $t \to +\infty$. Case 2 occurs below this region (coloured dark) and Case 3 above (coloured white), as we will see shortly.

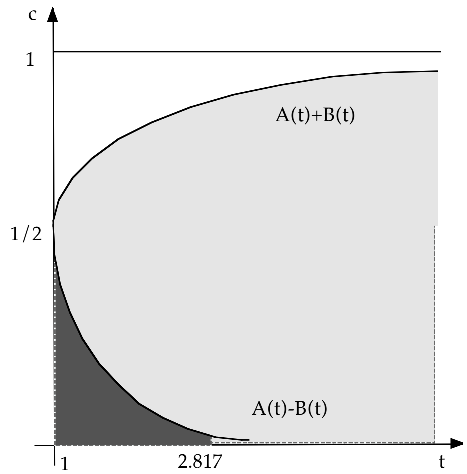

Fig. 1. The region for Planckett's example.

As far as the profit function is concerned, a simple substitution gives

$$\Pi(p^*(t),t) = \frac{(A+B-c)^2}{4B} = \frac{1}{4}\frac{[c(t-1)^2 - t(t-\ln t - 1)]^2}{(t-1)^2(t^2-2t\ln t - 1)}$$

• Case 2 ($c \le p^\circ < A(t)$). Working out the restriction, we find $c \le A+B$ and $c < A-B$, which can be reduced to $c < A-B$. For the profit function, a simple substitution gives

$$\Pi(p^*(t),t) = A(t) - c = \frac{1 + t\ln t - t}{(t-1)^2} - c$$

Since $A(t)$ is decreasing in t, in Case 2 the optimal choice for t would be $t^* = 1$ yielding an optimal profit (take limits to compute it) $\Pi(p^*(1),2) = (1/2) - c$. In Case 2, we have everybody buying: hence, it is preferable to have a noninformative signal. Note that, when Case 2 occurs, $c < A(t) \le 1/2$: this case can only occur when the cost c is not too high and requires the seller to charge a price $p^*(t) = A(t) \le 1/2$ which is not higher than what he could get with no information.

• Case 3 ($p^\circ < \min\{c, A(t)\}$ or $p^\circ > A(t)+B(t)$). Working out the restrictions, we find that the only binding constraint is $c > A+B$. For the profit function, no sale occurs and thus $\Pi(p^*(t),t) = 0$.

Summarizing, we can take into account the constraints as described above and give the optimized profit as

$$\Pi(p^*(t),t) = \begin{cases} \dfrac{1}{4}\dfrac{[c(t-1)^2 - t(t-\ln t - 1)]^2}{(t-1)^2(t^2-2t\ln t - 1)} & \text{if } A-B \le c \le A+B \\[2ex] \dfrac{1+t\ln t - t}{(t-1)^2} - c & \text{if } c < A-B \\[2ex] 0 \text{ (no sale)} & \text{if } c > A+B \end{cases}$$

or, alternatively,

$$\Pi(p^*(t),t) = \begin{cases} \dfrac{(A+B-c)^2}{4B} & \text{if } A-B \le c \le A+B \\[2ex] A(t) - c & \text{if } c < A-B \\[2ex] 0 \text{ (no sale)} & \text{if } c > A+B \end{cases}$$

Proposition 8. *The optimized profit function $\Pi(p^*(t),t)$ is quasiconvex in t.*

Proof. (Sketch) The function $\Pi(p^*(t),t)$ is continuous. Moreover, it is concave increasing in Case 1, convex decreasing in Case 2 and constant in Case 3. Look at Figure 1. For a given value of c, as t increases, we move from Region 3 (white) to Region 1 (grey) when c is high, or from Region 2 (dark) to Region 1 (grey) when c is low, or remain always in Region 1 (grey) when $c = 1/2$.

In the first case, we study

$$\Pi(p^*(t),t) = \frac{(A+B-c)^2}{4B}.$$

The total derivative is

$$\frac{\mathrm{d}\Pi}{\mathrm{d}t} = \frac{4(A+B-c)}{16B^2}\left[2B(A'+B') - B'(A+B-c)^2\right]$$

whose sign depends only on the bracketed term because the fraction always positive. Expliciting the term in square brackets, we find

$$[2B(A'+B') - \ldots] = -2\frac{(t\ln t + \ln t - 2t + 2)\left[t\ln t + 1 - t - c(t-1)^2\right]}{(t-1)^5}$$

whose sign is the opposite of the square-bracketed term in the numerator (the first item on the numerator and the denominator are always positive for $t > 1$). Thus, we only need to consider $\left[t\ln t + 1 - t - c(t-1)^2\right]$, which is positive if and only if

$$A(t) = \frac{t\ln t + 1 - t}{(t-1)^2} \ge c$$

It follows that $\Pi(p^*(t), t)$ is increasing (respectively, decreasing) if and only if $A(t_) \leq (\geq) c$. Recall that $A(t)$ is decreasing and its maximum is $1/2$. Therefore, for $c \geq 1/2$, $A(t) \leq c$ and $\Pi(p^*(t), t)$ is increasing. Instead, for $c \leq 1/2$, $A(t)$ is initially greater than c and then lower: thus $\Pi(p^*(t), t)$ is at first decreasing and then increasing.

In the second case, $\Pi(p^*(t), t)$ is convex decreasing first and then concave increasing: the profit is quasiconvex and the optimal choice is $t = 1$ or $t \to +\infty$ depending on the value of $\Pi(p^*(t), t)$ at these two points.

In the third (limit) case, the function is always concave increasing and again the optimal choice is $t \to +\infty$, where the profit is $1/2$.

Pasting together the three cases together, the function $\Pi(p^*(t), t)$ is shown to be quasiconvex in t and we are done.

References

1. Joe, H.: Multivariate Models and Dependence Concepts. Chapman and Hall, London (1997)
2. Johnson, J.P. and Myatt, D.P.: On the simple economics of advertising, marketing, and product design. Discussion Paper, Oxford University **185** (2004) November
3. Lehmann, E: Some concepts of dependence. Annals of Mathematical Statistics **37** (1966) 1137–1153
4. Lehmann, E.: Comparing location experiments. Annals of Statistics **16** (1988) 521–533
5. Lewis, T.R. and Sappington, D.E.M.: Supplying information to facilitate price discrimination. International Economic Review **35** (1994) 309–327
6. Milgrom, P. and Shannon, C.: Monotone comparative statics. Econometrica **62** (1994) 157–180
7. Moscarini, G. and Ottaviani, M.: Price competition for an informed buyer. Journal of Economic Theory **101** (2001) 457–493
8. Nelsen, R.B.: An Introduction to Copulas. Springer, New York (1999)
9. Persico, N.: Information acquisition in auctions. Econometrica **68** (2000) 135–148
10. Yanagimoto, Y. and Okamoto, M.: Partial orderings of permutations and monotonicity of a rank correlation statistic. Annals of the Institute of Statistical Mathematics **21** (1969) 489–506