<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in *Journal of Economics*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Manduchi, A. (2013)

Non-neutral information costs with match-value uncertainty

Journal of Economics, 109(1): 1-25

<https://doi.org/10.1007/s00712-012-0283-7>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:hj:diva-18345>

NON-NEUTRAL INFORMATION COSTS WITH MATCH-VALUE UNCERTAINTY

AGOSTINO MANDUCHI

Abstract. This paper investigates a model featuring a monopolist seller and a buyer with an uncertain valuation for the seller's product. The seller chooses an information system which allows the buyer to receive a private signal, potentially correlated with her valuation. No restrictions are imposed on the conditional distributions of the signal; the cost of the information system is proportional to its precision, measured by the mutual information between the distributions of the buyer's valuation and the signal. In equilibrium, the information system trades off the information cost against the losses deriving from a probability of trade that is either "too high," or "too low" - depending on the relative weight of the expected losses resulting from errors of the two types - and sends "non-neutral" signals, typically. Thus, in general, the probability of a correct signal depends on the buyer's actual valuation, and the probability of trade differs from the probability of a valuation exceeding the cost of production. The expected total surplus generated by the exchange is maximized, in equilibrium.

Keywords Match-value, Information provision, Mutual information, Bayesian learning

JEL Classification D42, D83, L10

1. Introduction

1.1. Motivation and results. Consumers and firms are often uncertain about the extent to which the products available for purchase can match their preferences, or the requirements of their production processes. Hence, providing information that can reduce such uncertainty is typically the aim of a substantial part of the sellers' marketing activities. Concretely, the sellers may provide detailed descriptions of their products, clarify how the buyers can benefit from them, and/or allow the buyers to physically inspect the products, or to run preliminary tests. This paper investigates a model of these activities where the seller chooses the conditional probability distributions characterizing an *information system*. The information system allows the buyer to receive a private signal, potentially correlated with her valuation. In the spirit of the literature on "rational inattention" - see [Sims \(2003\)](#) - the *mutual information* of the information system is used to measure the "quantity of information" (or information processing capacity) that the seller invests to make the conditional distributions more responsive to the buyer's valuation, and to thereby make more precise information available. No restrictions are imposed on the conditional distributions of the signal, and any given quantity of information is generally associated with a continuum of different information systems.

The paper establishes two related results:

- (1) The equilibrium information system is noisy, and generally "non-neutral," in the sense that (a) the probability of a correct signal differs across the buyer's possible valuations, and (b) the probability of trade differs from the probability of a valuation exceeding the cost of production.

Date: April 13, 2012.

Very detailed and valuable comments from two anonymous referees are gratefully acknowledged. I am also thankful to Tobias Dahlström, Per Davidsson, Anna Jenkins, seminar participants at the University of Gothenborg, and participants in the Workshop held during the 2011 Arne Ryde Lectures, at the University of Lund. The usual caveat applies.

- (2) The value of the optimal information system, for the seller, is a concave function of the quantity of information provided.

In equilibrium, if the expected surplus under the default probability measure is either very high or very low (and negative), the seller provides no information, and exchange respectively takes place either with probability 1, at the price justified by the prior beliefs, or with probability 0. For intermediate values of the surplus, the seller provides noisy information and chooses the highest price that the buyer is willing to accept if she observes a favorable realization of the signal.

If the seller does provide information, the general conditional distributions of the signal allow her to trade off the probability of a favorable signal against the increase in the valuation of the product, if the buyer observes the same signal, *even* with a constant quantity of information. Depending on whether the default expected valuation is relatively high or relatively low, compared to the cost of production, the seller's priority can be either a high probability of trade, or a large increase in the buyer's valuation, upon observation of a favorable signal. Hence, in the former case, the seller tends to be "bold," namely to choose a large unconditional probability of a favorable signal, and pays special attention to limiting the probability of an unfavorable signal being sent to a high valuation-buyer. In the latter case, the seller tends to be "cautious," namely to choose a small unconditional probability of a favorable signal, and is especially careful to limiting the probability of a favorable signal being sent to a low valuation-buyer, to boost the effect of such a signal on the buyer's willingness to pay. The downside, in the two cases, is respectively represented by a diluted impact of a favorable signal on the buyer's willingness to pay, and by the small number of positive recommendations issued.

The paper suggests that non-neutral information provision can be an equilibrium phenomenon, and can in fact be consistent with maximization of the expected surplus from trade - as it is the case in the model studied in the paper - even if the seller is effectively committed to the informational strategy announced, by assumption. In general, commitment could be guaranteed by an outside information provider, with a reputational concern. In the presence of incentives to overstate the probability of a high valuation, the seller's credibility would be more heavily strained in the case of products that are relatively unlikely to be a good match for the buyer. On the other hand, the same incentives may even be functional, within limits, for the sellers of products that are relatively likely to match the buyer's preferences. Thus, resorting to a third-party information provider might be a more useful option for sellers facing situations of the former type. This implication can complement the implication of [Lewis and Sappington \(1994, p. 320\)](#) that, if we consider products with a relatively high cost of production, we should expect the sellers to invest substantial amounts of resources to explain the characteristics and the uses of the products that they sell.

The investigation of the endogenous commitment opportunities which could be created by the observability of the seller's record as an information provider could be an interesting topic for future research. The information on record could be used either by the potential buyers, in a setting with repeated trades, and/or by a supervising public authority with the power to levy penalties and to grant rewards. The non-neutrality of the optimal information system, identified in the paper, highlights the potential importance of the consistency with the announced informational strategy as a criterion to evaluate the information provided.

1.2. Related literature. The model builds on [Lewis and Sappington \(1994\)](#), where the information systems available to the seller are subject to restrictions that make it possible to order them linearly by their informativeness. The choice of the information system has no direct payoff consequences, such as a cost, but only indirect consequences: More precise information increases the dispersion of the valuations corresponding to the different buyer types. Hence, it increases the prices that some buyer types are willing

to pay, but can also create informational rents. In a number of settings studied in the paper - including a model with two buyer types, providing the closest counterpart to the model of the present paper - the seller either refrains from providing information, or provides as much information as possible, in equilibrium.

Also, the adoption of a relatively noisy information system can lower the seller's profit. Essentially, more precise information lowers the price(s) that a buyer who receives unfavorable information is willing to accept, and a small improvement in the precision may be insufficient to make a price that is only accepted upon receipt of favorable information a part of the seller's optimal strategy. As [Lewis and Sappington \(1994, p. 317, footnote 16\)](#) observe, this result is similar to the non-concavity result established by [Radner and Stiglitz \(1984\)](#) in a setting with a single strategic player and costly information. By contrast, in the present model - as I mentioned in Subsection 1.1 above, and as I explain in more detail in Subsection 3.2 below - the value of information is a concave function of quantity. The different features of the two models essentially depend on the fact that our seller - unlike the seller in [Lewis and Sappington \(1994\)](#) - can choose among different information systems even for a *given* quantity of information. This situation is reminiscent of the statement in [Chade and Schlee \(2002, p. 446\)](#) that "the Radner-Stiglitz nonconcavity emerges only by severely constraining the set of information structures available to decision makers."

[Saak \(2006\)](#) shows that in a setting with unit demand, a seller with access to costless, general information structures would only allow the buyer to know if her match-value is greater or smaller than the cost of production. [Johnson and Myatt \(2006\)](#) consider restrictions under which the effects of different informational strategies (and marketing strategies in general) can be classified with reference to "rotations" and "shifts" of the demand function. Essentially, rotations and shifts respectively correspond to the provision of information about idiosyncratic match-values and of "hype" information, related to the availability and the quality of the product. Similarly to the case in [Lewis and Sappington \(1994\)](#), in the absence of cost differences, the seller typically chooses extreme informational strategies, in equilibrium. In [Anderson and Renault \(2006\)](#), a cost that the buyers must bear to visit the seller's shop creates a potential hold-up problem, and can thereby induce the seller to provide only a limited amount of information about the match-value, and possibly also to make a price commitment. Related problems are considered in [Villas-Boas \(2004\)](#), in connection with the seller's choice of the optimal product line, and in [Bar-Isaac et al. \(2010\)](#), where the seller can affect the transaction cost faced by the buyers. [Board \(2009\)](#) and [Ganuza \(2004\)](#) consider auctions where the auctioneer and the buyers respectively have private information about the object being sold and the own willingness to pay for the different possible types of the object. Information disclosure can soften competition, via its effect on the dispersion of the valuations. With relatively few bidders, the seller may thus release an amount of information smaller than the amount required to maximize the expected total surplus; the gap tends to disappear as the number of bidders increases. [Grossman and Shapiro \(1984\)](#), [Hamilton \(2009\)](#), and [Meurer and Stahl \(1994\)](#) consider the effects of competition on information provision in markets with horizontal differentiation.

In models of *ex-ante* vertical differentiation, the buyers' expectation that a monopolist would only withhold unfavorable information can lead to full disclosure, if the cost is negligible - see [Grossman \(1981\)](#), [Grossman and Hart \(1980\)](#) and [Milgrom \(1981\)](#). On the other hand, disclosure of costly information is only beneficial for the sellers of products with quality above an endogenous threshold ([Jovanovic, 1982](#)). The effects of competition on the sellers' incentives to disclose quality are investigated in [Milgrom and Roberts \(1986\)](#), [O. Board \(2009\)](#), and [Levin, Peck and Ye \(2009\)](#), among others. The literature on quality disclosure and certification is surveyed in [Dranove and Jin \(2010\)](#). [Sun \(2009\)](#) considers disclosure to heterogeneous buyers who are uncertain both about

the quality of a product and about their valuations for it. The equilibrium may feature a trade-off between the information of the two types available to the agents. Koessler and Renault (2012) provide a necessary and sufficient condition under which full disclosure is the unique equilibrium outcome in a model that also features twofold *ex-ante* uncertainty. The information-theoretic criterion used in the present paper to assign an explicit cost to general information systems could be used to explore interesting extensions of these papers, in future research.

Suen (2004) and Burke (2008) consider settings where, given the prior beliefs and the payoffs of the agents, asymmetric error probabilities make information more valuable, as in the present paper. In Suen (2004), some information providers transform a continuous signal into a binary signal by using exogenous, heterogeneous thresholds. The providers' reports differ in a (random, but) systematic way, and a decision-maker optimally chooses a provider whose report is relatively likely to confirm her prior beliefs. This choice is justified by the fact that only realizations of the binary signal that both the decision-maker and the provider regard as unlikely, corresponding to extreme realizations of the underlying signal, can possibly induce the decision-maker to change the default decision, and can thereby make the report valuable, *ex-ante*. In Burke (2008), the information providers may "bias" their reports, by choosing an asymmetric distribution of a given total probability of a correct signal across states; competition between two providers exacerbates the distortions, compared to the monopoly benchmark, essentially by forcing each provider to target a specific segment of the audience. The main feature of the present paper, in comparison with Burke (2008) and Suen (2004), is the use of a criterion rooted in the information-theoretical literature, such as mutual information, to assign a cost to general information systems. This modeling strategy allows me to avoid restrictions that can otherwise be required to reconcile tractability and comparability, but that can be difficult to justify, in general.

A related paper by Ganuza and Penalva (2010) investigates a number of measures that - similarly to mutual information - focus on the variability of the conditional distributions if the prior beliefs are held fixed. A measure called *supermodular precision*, according to which a signal ϕ is more informative than another signal ϕ' if the difference between any two quantiles of the distribution of the conditional expectation is smaller under ϕ' than under ϕ , is used to investigate costly information provision by a seller auctioning an object. In the setting of the present paper, each information system features at most two distinct realizations of the willingness to pay for the product, *ex-post*. Hence, two non-degenerate information systems can only possibly be ranked if they feature an identical probability of a favorable realization of the signal; otherwise, for each distribution, we can always find a pair of quantiles that coincide under that distribution, but not under the other one. A comparison between the information measures in Ganuza and Penalva (2010) and mutual information, in general settings, could be an interesting topic for future research. Mutual information is considered by Cabrales, Gossner and Serrano (2012) in connection with the problem of an investor choosing a portfolio of risky assets. The assets' prices and the returns are such that no arbitrage opportunities exist under the prior probability measure. Under some restrictions on preferences, the investor's willingness to pay for a signal is necessarily measured by the mutual information between the signal and the state of the world.

1.3. Plan of the paper. The model is presented in Section 2. The equilibrium and the non-neutrality of the information system are respectively characterized in Section 3 and in Section 4. Section 5 contains some concluding remarks, and indicates possible topics for future research. All proofs are in the Appendix. The numerical examples on which the figures are based were developed with the help of *wxMaxima*, which was also used to verify the calculations; the files are available from the Author upon request.

2. The model

There are a monopolist (producer-)seller of an indivisible product, and a (potential) buyer, who is willing to purchase at most one unit of the product. Both agents are risk-neutral, and maximize their expected payoffs. Production can take place upon demand, at the cost $c \in \mathbb{R}_{++}$. The buyer's type θ is equal to H or to L with respective probabilities q and $1 - q$, where $q \in (0, 1)$ is given; the valuations v_H and v_L are real numbers such that $\max\{v_L, 0\} < c < v_H$.¹ The results of the paper would still hold if we replaced the buyer by a $[0, 1]$ continuum of *ex-ante* identical buyers, with the subsets of the H - and the L -buyers having respective measures q and $1 - q$, and endowed the seller with a production technology featuring constant returns to scale.

As in [Lewis and Sappington \(1994\)](#), the distribution of the buyer's type is common knowledge. For the buyer, imperfect knowledge of the own type could be due to the lack of familiarity with the product, which could in turn follow from the novelty of either the product or the situations in which it could be used. Additionally, the buyer observes a private signal $s \in \{h, l\}$, issued by an information system chosen by the seller. For example, the seller could allow the buyer to test the product, at no charge, or provide free leaflets and presentations allowing the buyer to receive detailed information. Formally, I model an admissible information system as a pair $(a_H, a_L) \in [0, 1]^2$, which defines the conditional distributions of s . For each buyer type $\theta \in \{H, L\}$, we have:

$$s = \begin{cases} h, & \text{with probability } a_\theta, \\ l, & \text{with probability } 1 - a_\theta. \end{cases}$$

I restrict attention to information systems featuring $a_H \geq a_L$, so that the posterior probability of a type H -buyer is (weakly) greater or smaller than the prior probability q , respectively depending on whether $s = h$ or $s = l$.² This assumption entails no loss of generality; if h and l were systematically swapped, the results of the paper would still hold.

The cost of the information system, borne by the seller, is proportional to the mutual information between the distributions of the buyer type and the realization of the signal. For notational convenience, given any $x \geq 0$, I set

$$r(x) = \begin{cases} 0, & \text{if } x = 0, \\ -x \log(x) & \text{otherwise.} \end{cases} \quad (1)$$

Also, I use

$$A = qa_H + (1 - q)a_L \quad (2)$$

to denote the unconditional probability of $s = h$. The quantity of information provided by our “double binary” information system is then

$$I(a_H, a_L) = r(A) + r(1 - A) - q(r(a_H) + r(1 - a_H)) - (1 - q)(r(a_L) + r(1 - a_L)); \quad (3)$$

see [Cover and Thomas \(1991\)](#).³ The argument of $r(x)$, in (3), is always the probability of some event; $r(x)$ can be interpreted as the product of the same probability and the

¹If $c \leq v_L$, the seller would never provide information; formally, (12) below would be maximized with $a_H = a_L = 1$, and the buyer would purchase the product with probability 1. Negative values of v_L , rationalized for example in terms of a cost that the buyer must bear to dispose of an unsuitable product, are allowed.

²Thus, we are only considering pairs (a_H, a_L) in a proper subset of $[0, 1]^2$; this restriction is not explicitly formalized, to avoid unnecessary notational complications.

³The notation used here, emphasizing the seller's choice variables a_H and a_L , differs from the standard notation used in the information theory literature - see, for example, [Cover and Thomas \(1991\)](#) and [MacKay \(2003\)](#) - where the symbols denoting the random variables would typically be used as the arguments of I . Also, the basis of the logarithm in (1) is e , *not* 2, the more common basis in information theory. Thus, information is measured in *nats*, rather than in *bits*. The latter choice simplifies the expressions for the derivatives, and

Shannon information content of the event, $-\log(x)$ - see MacKay (2003). $I(a_H, a_L)$ is the expected “quantity of information” that knowledge of the realization of s provides about the realization of θ , measured in terms of Shannon’s entropy - see Shannon (1948).⁴ Information systems may be either *trivial* or *non-trivial*, depending on whether the signal is pure noise, or whether it conveys relevant information to the buyer, namely on whether $a_H = a_L$ or $a_H > a_L$. The cost of a generic information system (a_H, a_L) is equal to $kI(a_H, a_L)$, where $k \in \mathbb{R}_{++}$ is the given unit cost. Alternatively, we could specify a capacity constraint $\bar{I} \in (0, I(1, 0))$, as in Sims (2003), and recover k as the “shadow cost” of the quantity of information provided.

Any trivial information system provides a quantity of information equal to 0. At the opposite extreme, with two buyer types, the system featuring a one-to-one correspondence between the buyer type and the signal provides a quantity of information equal to the entropy of the distribution of the types, namely to $I(1, 0) = r(q) + r(1 - q)$.⁵ Lewis and Sappington (1994) generally consider a signal which is either informative, albeit possibly noisy, or uninformative. If we use β to denote the (type-independent) probability of an informative signal, characterized by a generic pair (a_H, a_L) , a system of this type would provide a quantity of information equal to $\beta I(a_H, a_L)$, by additivity of entropy. A similar situation occurs in Johnson and Myatt (2006). Lewis and Sappington (1994) also consider a model with two buyer types, in which the seller chooses the symmetric probability of error of a signal in a continuum of possible values. This model provides the closest point of reference for the model of the present paper; the information provided by it, with generic parameter $a \in [\frac{1}{2}, 1]$ is equal to $I(a, 1 - a)$. In general, any feasible quantity of information, except $I(1, 0)$, corresponds to a continuum of different information systems.

In information theory, mutual information is generally used as a measure of the capacity of an information channel - both the capacity that would be necessary for the error-free transmission of a set of potential messages - and the actual capacity available. In the present paper, reference to mutual information can be rationalized by thinking that a larger investment in media exposure and/or sales personnel allows the seller to send a greater number of messages to the buyer. Such investment can reduce the probability of an incorrect classification of the type, due to the buyer’s limited attention and/or information-processing skills. With a continuum of buyers, endowed with heterogeneous heuristics, a broader set of “styles” could also allow a number of buyers representing a larger fraction of the population to interpret the message correctly.

The buyer is assumed to be fully attentive. The qualitative results of the paper are robust to the possibility of wrong responses to the signal observed, as long as the probability of such responses is “not strongly affected” by the structure of the information system. In the interpretation of the model with a continuum of buyers, a more careful analysis of the informational problem and a greater amount of resources devoted to it could be justified, for the seller, by the larger payoff at stake. An explicit attention cost borne by the buyer would require the seller to commit to the price, to avoid the no-trade result of Diamond (1971): Without commitment, the *ex-post* optimal price would in fact leave the buyer with a negative net surplus, if she chose to participate in the market.

has no impact on the results of the paper. e could be replaced by any basis $g \in \mathbb{R}_{++}$, provided k were correspondingly multiplied by $\log(g)$.

⁴The mutual information is also equal to the difference between the entropy of the hypothetical joint distribution of θ and s corresponding to the marginal distributions, if the two variables were independent, and the entropy of the actual joint distribution; see Cover and Thomas (1991).

⁵More generally, the quantity of information provided under the equilibrium strategy in Saak (2006), where information is costless, is equal to $r(Q) + r(1 - Q)$, where Q is the probability of a type with valuation no lower than c . If the valuation of a buyer type is equal to c , the model admits multiple equilibria, corresponding to multiple quantities of information.

Bar-Isaac et al. (2010) investigate a model in which the seller's marketing strategy determines the cost that the buyers must bear, if they choose to gather information about the product before making their purchase decision. An extension of the present model along similar lines is a potentially interesting topic for future research.

The “double-binary” structure of the communication channel allows me to obtain analytical solutions for the endogenous variables. With more than two buyer types, equilibrium and optimal (active) information provision would still be achieved with a binary signal, whose realizations are interpreted as recommendations to purchase the product, or not, as in the multi-type setting with costless information in Saak (2006). As in the present paper, the probability of a favorable realization of the signal, for each buyer type, would be an increasing function of the gain from trade - possibly only weakly increasing, if the equilibrium information system is trivial.

The equilibrium notion is *Perfect Bayesian Equilibrium* (“PBE”) - see Fudenberg and Tirole (1991). The timing is as follows:

- (1) The seller chooses the information system (a_H, a_L) , and the price $p \in \mathbb{R}$, and makes a take-it-or-leave-it offer at such price.
- (2) The buyer observes (a_H, a_L) and the realization of the signal s , updates her beliefs about her type, by using Bayes' rule, and decides whether to purchase the product, or not.

In principle, the seller could mislead the buyer by announcing $(a_H, a_L) = (1, 0)$, and choosing $(a_H, a_L) = (1, 1)$ instead. However, in the present model, observability of (a_H, a_L) effectively guarantees the seller's commitment to the information system chosen, in the seller's own interest. As I already noted in 1.1, the possibility of endogenous commitment could represent an interesting topic for future research. As the signal is privately observed by the buyer, no commitment problem arises in connection with the price.

As the seller has no informational advantage over the buyer, the choice of information systems at odds with the equilibrium system cannot possibly reveal any private information, and the buyer's beliefs in a PBE are necessarily formed by using Bayes' rule, both on and off the equilibrium path. For any given system chosen by the seller, I use w_s to denote the buyer's willingness to pay conditional on any realization of the signal that has a positive probability, formed in accordance with Bayes' rule:

$$w_s = \begin{cases} v_L + \frac{q a_H (v_H - v_L)}{A}, & \text{if } s = h, \\ v_L + \frac{q(1-a_H)(v_H - v_L)}{1-A}, & \text{if } s = l, \end{cases} \quad (4)$$

where A is defined in (2). It is also convenient to use

$$W = v_L + q(v_H - v_L) \quad (5)$$

to denote the buyer's default expected valuation - which is also the expected valuation conditional on a signal issued by a trivial information system.

The buyer's acceptance strategy is a pair (b_h, b_l) ; for each $s \in \{h, l\}$, b_s is a function from $[0, 1]^2 \times \mathbb{R}_+$ into $[0, 1]$, and expresses the probability of the product being purchased upon observation of s , given the price p , the information system (a_h, a_l) , and the Bayesian posterior in (4) - see (8) below. The seller's expected profit and the seller's problem are

$$\begin{aligned} \pi(a_H, a_L, p; b_h, b_l) = \\ ((q a_H + (1 - q) a_L) b_h + (q(1 - a_H) + (1 - q)(1 - a_L)) b_l) (p - c) - k I(a_H, a_L) \end{aligned} \quad (6)$$

and

$$\max_{(a_H, a_L, p) \in [0, 1]^2 \times \mathbb{R}} \pi(a_H, a_L, p; b_h, b_l). \quad (7)$$

The actual requirements for equilibrium are stated in Definition 1.

Definition 1. *An equilibrium is an information system (a_H^*, a_L^*) and a price p^* , for the seller, and an acceptance strategy (b_h^*, b_l^*) , for the buyer, such that:*

(1) *The acceptance strategy satisfies*

$$b_S^*(a_H, a_L, p) = \begin{cases} 1, & \text{if } p^* < w_S^*, \\ b, \text{ for an arbitrary } b \in [0, 1], & \text{if } p^* = w_S^*, \\ 0, & \text{if } p^* > w_S^*. \end{cases} \quad (8)$$

where S is any realization of s with a positive probability, given (a_H^*, a_L^*) and p^* .

(2) *The information system and the price solve (7), given the acceptance strategy in Part 1.*

Thus, in a PBE, the buyer accepts or rejects the seller's offer with probability 1 if the price is respectively above or below her willingness to pay in (4); both acceptance and rejection can be chosen with a positive probability if equality holds. Asterisks are used to denote the equilibrium values of all endogenous variables.

3. Equilibrium

3.1. Preliminary results. Lemma 1 essentially states that trade takes place either with probability 1, at a price equal to the buyer's valuation, or with probability 0, depending on whether the valuation justified by the realization of the signal is greater or smaller than the cost of production. Furthermore, if the seller chooses to provide information, the probability of trade necessarily depends on the realization of the signal.

Lemma 1.

- (1) *Let S denote any realization of the signal observed with a positive probability in equilibrium. Then with a trivial information system, $c < W$ and $c > W$, for W defined in (5), respectively imply $p^* = W$ and $b_S^*(a_H^*, a_L^*, p^*) = 1$, and $p^* \geq W$ and $b_S^*(a_H^*, a_L^*, p^*) = 0$. If $c = W$, then we can have either $p^* \geq W$ and $b_S^*(a_H^*, a_L^*, p^*) = 0$, or also $p^* = W$ and $b_S^*(a_H^*, a_L^*, p^*) \in (0, 1]$.*
- (2) *With a non-trivial information system, we have $p^* = w_h^*$, $b_h^*(a_H^*, a_L^*, p^*) = 1$, and $b_l^*(a_H^*, a_L^*, p^*) = 0$.*

Lemma 1 leaves open the possibility of multiple values of the probabilities of acceptance $b_h^*(a_H^*, a_L^*, p^*)$, $b_l^*(a_H^*, a_L^*, p^*)$ and/or the price $p^* \geq c$ only if $W \leq c$, and the seller provides no information. This multiplicity is however irrelevant, as both the buyer's and the seller's expected payoffs are equal to 0 in this scenario. To avoid unnecessary notational complications, I therefore assume that a seller who provides no information sets $(a_H^*, a_L^*) = (\hat{a}_H, \hat{a}_L)$, where

$$(\hat{a}_H, \hat{a}_L) = \begin{cases} (1, 1), & \text{if } c < W, \\ (a, a), \text{ for an arbitrary } a \in [0, 1], & \text{if } c = W, \\ (0, 0), & \text{if } c > W, \end{cases} \quad (9)$$

and

$$p^* = \max \{w_h^*, c\}, \quad (10)$$

and that the buyer's acceptance probabilities, in the face of the seller's equilibrium strategy, are

$$(b_h^*(a_H^*, a_L^*, p^*), b_l^*(a_H^*, a_L^*, p^*)) = (1, 0). \quad (11)$$

By using (9), (10) and (11), we can rewrite the expected profit (6) and the seller's problem (7), for operational purposes, as

$$\Pi(a_H, a_L) = qa_H(v_H - c) + (1 - q)a_L(v_L - c) - kI(a_H, a_L) \quad (12)$$

and as

$$\max_{(a_H, a_L) \in [0, 1]^2} \Pi(a_H, a_L). \quad (13)$$

Remark 1. *In a setting with unit demand, the seller's problem (13) coincides with the problem faced by a hypothetical planner maximizing the total expected surplus for any number $n \geq 2$ of buyer types.*

To justify Remark 1, let $v_\theta \in \mathbb{R}^n$ denote the willingness to pay of type $\theta \in \{1, \dots, n\}$, with $v_i < c < v_j$ for some indices i and j , and let $\{q_1, \dots, q_n\}$ denote the respective probabilities. A result corresponding to Lemma 1 above, echoing the analysis of the case with costless information systems in Saak (2006), would allow us to focus on the case in which the system issues either a favorable signal h , with conditional probabilities $(a_1, \dots, a_n) \in [0, 1]^n$, or an unfavorable signal l , and the price is equal to the buyer's valuation if $s = h$. Using n as a superscript to denote the expected profit and the mutual information in the modified setting, with a convention corresponding to that introduced in (9), (10) and (11), and observing that $w_h = \frac{\sum_{\theta=1}^n q_\theta a_\theta v_\theta}{\sum_{\theta=1}^n q_\theta a_\theta}$, in a PBE, we can then establish equality between the expected profit and the expected surplus by noting that

$$\begin{aligned} \Pi^n(a_1, \dots, a_n) &= \left(\sum_{\theta=1}^n q_\theta a_\theta \right) \cdot (w_h - c) - kI^n(a_1, \dots, a_n) \\ &= \sum_{\theta=1}^n q_\theta a_\theta (v_\theta - c) - kI^n(a_1, \dots, a_n). \end{aligned} \quad (14)$$

Lemma 2 states that $\Pi(a_H, a_L)$, in (12), is a concave function of the probability pair characterizing the information system - and a *strictly* concave function, if we consider non-trivial information systems.

Lemma 2. *$\Pi(a_H, a_L)$ is concave. In particular, if (a'_H, a'_L) and (a''_H, a''_L) are two distinct information systems featuring either $a'_H \neq a'_L$, and/or $a''_H \neq a''_L$, then for any $\lambda \in (0, 1)$ and the respective information system $(a^\lambda_H, a^\lambda_L) = \lambda(a'_H, a'_L) + (1 - \lambda)(a''_H, a''_L)$, we have the strict inequality*

$$\lambda \Pi(a'_H, a'_L) + (1 - \lambda) \Pi(a''_H, a''_L) < \Pi(a^\lambda_H, a^\lambda_L). \quad (15)$$

Lemma 2 implies that non-unique solutions to (13) would necessarily feature trivial information systems.

The first order conditions for (13) - obtained by setting (A.1) and (A.2) in Part A.3 of the Appendix equal to 0 - are both necessary and sufficient for a solution at which the seller does in fact provide some information. To formalize this statement, we must confront the fact that as $r(x)$, defined in (1), is not differentiable at $x = 0$, the first order conditions for the problem in (13) - obtained by setting (A.1) and (A.2), in A.3, equal to 0 - are only defined in the interior of $[0, 1]^2$. Lemma 3 rules out non-trivial information systems featuring a boundary value of either a_H and/or a_L , in equilibrium - essentially, any such information system is dominated by "interior" systems conveniently close to it.

Lemma 3. *Any information system that solves (13) and features $a_H > a_L$ is necessarily in the interior of $[0, 1]^2$.*

A corollary of Lemma 3 is that the optimal information system necessarily sends incorrect signals with a positive probability. Lemma 4 concludes that any non-trivial (and necessarily interior) information system that is a local solution to the seller's problem (13) is indeed "the" solution to (13).

Lemma 4. *Any local solution to (13) featuring $a_H > a_L$ is a unique, global solution.*

Lemma 3 and Lemma 4 justify the strategy used below, in Part 1 of Theorem 1, to establish existence of an equilibrium: We prioritize the search for interior values of (a_H, a_L) that solve (13), and consider possible boundary solutions - which can be located either at $(0, 0)$, or at $(1, 1)$ - only if no such values exist.

We can illustrate the characterization of the optimal non-trivial information systems with reference to the case with $n \geq 2$ buyer types, already mentioned in connection with Remark 1. Let $\mu_\theta(s)$ and $\rho_\theta = \frac{\mu_\theta(h)}{\mu_\theta(l)}$ denote the posterior probability of a buyer of type $\theta \in \{1, \dots, n\}$, formed upon observation of signal $s \in \{h, l\}$, and the likelihood ratio for each type. From the first order conditions for an interior maximum of $\Pi^n(a_1, \dots, a_n)$ in (14), for each type θ , we obtain

$$\rho_\theta = \frac{a_\theta}{1 - a_\theta} \frac{1 - A^n}{A^n} = e^{\frac{v_\theta - c}{k}}, \quad (16)$$

where $A^n = \sum_{i=1}^n q_i a_i$. Thus, with a constant unit cost of information, the likelihood ratios are independent of the prior probability measure. Essentially, the marginal effect of a change in a_θ on the seller's expected revenue, gross of the information cost, is equal to $q_\theta (v_\theta - c)$. The corresponding marginal cost, in absolute value, is equal to the logarithm of the likelihood ratio for the given type, multiplied by the probability of the type and the unit information cost; if $v_\theta < c$, the likelihood ratio is smaller than 1, and a greater precision and cost correspond to *smaller* values of a_θ . Greater (absolute) values of the logarithms of the likelihood ratios thus correspond to greater values of the marginal information cost, and must be justified by a greater marginal revenue. Given convexity of mutual information, an optimal interior strategy may also exist with a mildly concave information cost. A concave and a convex information cost would respectively make the likelihood ratios more and less responsive to the payoff consequences of the errors.

Considering the expressions for ρ_θ provided by (16), a system comprised of n equations of the form

$$\mu_\theta(h) - \rho_\theta \mu_\theta(l) = 0, \quad (17)$$

for $\theta \in \{1, \dots, n\}$, and two equations of the form

$$\sum_{\theta=1}^n \mu_\theta(s) = 1, \quad (18)$$

for $s \in \{h, l\}$, if $n = 2$, identifies the posterior probabilities corresponding to both types, for both realizations of the signal. The probabilities of a buyer of type H formed upon observation of $s = h$ and of $s = l$ are

$$\bar{q} = \frac{1 - e^{-\frac{c - v_L}{k}}}{1 - e^{-\frac{v_H - v_L}{k}}} \quad (19)$$

and

$$\underline{q} = \frac{e^{\frac{c - v_L}{k}} - 1}{e^{\frac{v_H - v_L}{k}} - 1}. \quad (20)$$

In the binary type case of the paper, the posterior probabilities are thus also independent of the prior probabilities, as far as the seller does provide information. Each unit increase in n introduces two new variables, namely the posterior probabilities $\mu_\theta(h)$ and $\mu_\theta(l)$ for the new type, one additional equation of the type (17), and no additional equations of the type (18). Hence, with more than two buyer types, different prior probability measures for which the seller does provide information can be associated with different posterior probabilities - and some results available suggest that this does indeed happen with three types.

Let q_χ denote the probability of a buyer of type H that makes the expected surplus generated by the exchange equal to 0:

$$q_\chi = \frac{c - v_L}{v_H - v_L}. \quad (21)$$

Lemma 5 confirms that $\underline{q}$, q_χ , and $\bar{q}$ do qualify as probabilities, and provides information about their relative positions.

Lemma 5. $\underline{q}$ and $\bar{q}$ are respectively strictly increasing and strictly decreasing in k , with $\lim_{k \rightarrow 0} \underline{q} = 0$, $\lim_{k \rightarrow 0} \bar{q} = 1$, and $\lim_{k \rightarrow \infty} \underline{q} = \lim_{k \rightarrow \infty} \bar{q} = q_\chi$.

3.2. Existence and comparative statics. In order to state the results on existence and comparative statics, in Theorem 1, let us define

$$\alpha_H = \frac{qe^{\frac{v_H}{k}} + (1-q)e^{\frac{v_L}{k}} - e^{\frac{c}{k}}}{q \left(e^{\frac{v_H}{k}} - e^{\frac{v_L}{k}} \right) \left(1 - e^{-\frac{v_H - c}{k}} \right)}, \quad (22)$$

$$\alpha_L = \frac{qe^{\frac{v_H}{k}} + (1-q)e^{\frac{v_L}{k}} - e^{\frac{c}{k}}}{(1-q) \left(e^{\frac{v_H}{k}} - e^{\frac{v_L}{k}} \right) \left(e^{\frac{c - v_L}{k}} - 1 \right)}. \quad (23)$$

Theorem 1.

(1) The equilibrium information system is:

$$(a_H^*, a_L^*) = \begin{cases} (0, 0), & \text{if } q \leq \underline{q}, \\ (\alpha_H, \alpha_L), & \text{if } q \in (\underline{q}, \bar{q}), \\ (1, 1), & \text{if } q \geq \bar{q}. \end{cases} \quad (24)$$

(2) If $q \in (\underline{q}, \bar{q})$, a_H^* and a_L^* are both increasing in q , v_H , and v_L , and decreasing in c . Furthermore, we have $\lim_{k \rightarrow 0} a_H^* = 1$, $\lim_{k \rightarrow 0} a_L^* = 0$ and

$$\lim_{k \rightarrow \infty} a_H^* = \lim_{k \rightarrow \infty} a_L^* = \begin{cases} 0, & \text{if } q \in (\underline{q}, q_\chi), \\ \frac{1}{2}, & \text{if } q = q_\chi, \\ 1, & \text{if } q \in (q_\chi, \bar{q}). \end{cases}$$

Thus, equilibrium is characterized by (10), (11), and (24), given (9). If q is above $\bar{q}$ or below $\underline{q}$, trade occurs with respective probabilities 1 and 0. If $q \in (\underline{q}, \bar{q})$, the seller does provide information, the buyer's decision depends on the signal observed by the buyer, and the probability of trade increases monotonically with q .

As we know from Lemma 5, q_χ is asymptotically “sandwiched” between $\underline{q}$ and $\bar{q}$, with the latter variables monotonically converging toward it from the values of 0 and 1, as k becomes arbitrarily large. Providing information is a very valuable opportunity if q is close to q_χ , namely if the total (and the seller's) expected surplus that would be generated (or destroyed) by an exchange, with a trivial information system, is relatively close to the no-trade value of 0. As was mentioned after the definition of the relevant part of the buyer's acceptance strategy in (8), in Subsection 3.1, the case of $W = c$ (and $q = q_\chi$) would be the only one potentially consistent with multiple, reciprocally consistent best responses, if the seller provided no information. However, if $q = q_\chi$, the seller does provide information for *any* value of k , and equilibrium is therefore unique for every admissible parametrization of the model - up to a permutation of h and l .

In Figure 1, the results in Part 1 of Theorem 1 are illustrated by a plot of the quantity of information provided, in equilibrium, for different values of k and q ; the values of the remaining parameters are $c = \frac{5}{9}$, $v_H = 1$, and $v_L = 0$. Essentially, as k increases, the interval of the values of q for which the seller provides information becomes more and more narrow, and the quantity of information provided, within the interval, decreases.

Concerning the comparative statics results in Part 2 of Theorem 1, the changes in a_H^* and a_L^* essentially reflect the opportunity to reduce the (relative) probability of the

Figure 1. The equilibrium (and optimal) quantity of information provided for different pairs (k, q) , given the values of the remaining parameters.

error which becomes relatively more costly. For example, an increase in v_H increases the expected revenue loss from an unfavorable realization of the signal occurring if the buyer is of type H , and therefore immediately makes an increase in a_H payoff-increasing. A higher value of a_H translates in turn into an increase in the unconditional probability of $s = h$, A , and thereby makes any given value of a_L more costly, on the margin. A (relatively small) increase in a_L is therefore also necessary to restore the seller's first order conditions. This account can easily be adjusted to cover the case of changes in v_L . Increases of q and decreases of c immediately increase the expected marginal revenues associated with larger values of *both* a_H and a_L - and therefore, again, of A . From the point of view of the informational cost, the common direction of the changes has a beneficial effect, which may be optimally exploited by increasing the information provided.

If the unit information cost k is small, α_H and α_L are close to 1 and to 0, respectively; if the same cost is large, *both* α_H and α_L are close to 1, or to 0, depending on whether $c < W$ or $c > W$ holds. As it turns out, the variable which has the same value both with a relatively small and with a relatively large unit cost first moves away from this value, and then moves back toward it, as the cost increases, while the other variable features a monotonic transition between the two limiting values.

This situation is illustrated in Part (a) of Figure 2. Both parts of the Figure refer to an example with $c = \frac{5}{9}$, $q = \frac{2}{5}$, $v_H = 1$, $v_L = 0$, and therefore $q < q_\chi$; hence, if the seller provided no information, no trade would take place. On the axes, I measures the conditional probabilities of a realization of the signal equal to h for a buyer of type H and of type L . Given the acceptance probabilities $(b_h^*, b_l^*) = (1, 0)$ in (II), these probabilities coincide with the conditional probabilities of trade. Each black segment is the locus of the information systems such that the profit gross of the information cost, $qa_H(v_H - c) + (1 - q)a_L(v_L - c)$, is equal to a given constant. Each dark-gray curve is the locus of the information systems whose cost $kI(a_H, a_L)$ is equal to a given constant. The light-gray curve is the “expansion path,” parametrized by different values of k in $\mathbb{R}_{++}$; notice that $(a_H, a_L) = (1, 0)$ is never an optimal information system, as we know from Lemma 3, and that $(a_H, a_L) = (0, 0)$ is the optimal information system for all values of k above a given threshold.

Remark 2. *Considering the results in Part 1 of Theorem 1, together with continuity of $I(a_H, a_L)$, we can conclude that any quantity of information strictly lower than the greatest admissible value $I(1, 0)$ can be associated with an optimal information system, for a suitable value of k .*

Thus, the present model does *not* share the non-concavity in the value of information featured by the model in Lewis and Sappington (1994). Concavity of the value of

Figure 2. A graphical representation of the seller's optimization problem. Part (a) illustrates the case with general error probabilities. Part (b) illustrates the case in which the symmetry restriction $a_H = 1 - a_L$ is imposed, and the set of the strategies corresponding to a non-negative expected profit is non-convex.

information reflects the fact that the set of the information systems corresponding to a non-negative expected profit under the requirement of $(b_h^*, b_l^*) = (1, 0)$, stated in (II), is convex for each value of the unit cost k , and always includes its unique member for large values of k - either $(0, 0)$ or $(1, 1)$, depending on whether $c > W$ or $c < W$. Furthermore, the sets corresponding to different values of k can be monotonically ordered by (weak) inclusion.

A different situation would arise if the seller could only choose an information system featuring symmetric error probabilities, namely $a_H = 1 - a_L$, as in the model with two buyer types in Lewis and Sappington (1994). The situation is illustrated in Part (b) of Figure 2, where the light-gray segment represents the whole set of the information systems that would allow the seller to improve her *gross* expected profit - rather than the actual expansion path, as in Part (a). Even the former set is at a positive distance from the set of the trivial information systems.⁶ Hence, each information system belonging to it necessarily features a strictly positive cost, and information systems in the upper part of the segment can also be ruled out, as the values of $qa_H(v_H - c) + (1 - q)a_L(v_L - c)$ corresponding to them, albeit positive, would be too small to cover their cost. Thus, values of I between 0 and some $\iota \in (0, I(1, 0))$ cannot possibly be associated with an equilibrium.

Coincidence between the seller's problem and the problem of maximizing the expected total surplus, also with more than two buyer types - noted in Remark 1, in 3.1 -

⁶To achieve a probability of trade equal to 0 with $(b_h^*, b_l^*) = (1, 0)$, we could drop the symmetry requirement if the information system is trivial, thereby effectively assuming that the system can simply be left unused. Alternatively, we could identify the trivial information system with the system featuring $(a_H, a_L) = (\frac{1}{2}, \frac{1}{2})$; with $p^* = c$, as required by (10), (II) should then be replaced by $(b_h^*, b_l^*) = (0, 0)$.

allows us to conclude that the expected total surplus generated by the exchange is maximized in equilibrium. Surplus maximization follows from the fact that the posted price allows the seller to appropriate the whole surplus, and to thereby fully internalize the additional payoffs created by the information provided; this situation is essentially due, in turn, to the assumptions of unit demand and of common knowledge of the buyer's prior beliefs. Any private information available to the buyer, besides the signal issued by the seller's information system, could create informational rents, and thereby drive a wedge between the incentives faced by the seller and by a surplus-maximizing planner.

4. Non-neutrality

In the absence of an established convention, I identify a *neutral* information system, alternatively, with:

- (1) An information system such that the conditional probability of a correct signal is identical across buyer types, namely that

$$a_H^* = 1 - a_L^*. \quad (25)$$

- (2) An information system such that the unconditional probability with which each signal is sent coincides with the probability of the buyer type ideally associated with the signal, namely that

$$A^* = q. \quad (26)$$

The requirement in (25) is imposed in the two type-version of the model in Lewis and Sappington (1994), mentioned in Section 2 above, and is equivalent to the requirement in Definition 1 in Burke (2008). To justify (26), notice that Lemma 1, (II), and Part 1 of Theorem 1 allow us to conclude that A^* - the equilibrium value of the unconditional probability of $s = h$, defined in (2) - coincides with the *ex-ante* probability of trade.

If we impose symmetric error probabilities, and set $a_\sigma = a_H^* = 1 - a_L^*$, the sign of

$$qa_\sigma + (1 - q)(1 - a_\sigma) - q = (1 - 2q)(1 - a_\sigma),$$

coincides with the sign of $1 - 2q$. Thus, with a noisy information system satisfying (25), the signal ideally associated with the more likely buyer type is sent with an unconditional probability that is lower than the probability of the buyer type, and (26) is necessarily violated, except if $q = \frac{1}{2}$.

In order to state the result of the present Section, in Theorem 2, let

$$q_{sym} = \begin{cases} \frac{1}{2}, & \text{if } c = \frac{v_H + v_L}{2}, \\ e^{\frac{v_H}{2k}} \left(e^{\frac{c}{k}} - e^{\frac{v_L}{k}} \right) \frac{e^{\frac{v_L}{2k}} \left(e^{\frac{v_H}{k}} - e^{\frac{c}{k}} \right) - e^{\frac{v_H}{2k}} \left(e^{\frac{c}{k}} - e^{\frac{v_L}{k}} \right)}{\left(e^{\frac{v_H}{k}} - e^{\frac{v_L}{k}} \right) \left(e^{\frac{v_H + v_L}{k}} - e^{\frac{2c}{k}} \right)}, & \text{otherwise,} \end{cases} \quad (27)$$

and

$$q_{iso} = \frac{e^{\frac{c}{k}} - e^{\frac{v_L}{k}}}{e^{\frac{v_H + v_L}{k} - c} - 2e^{\frac{v_L}{k}} + e^{\frac{c}{k}}}. \quad (28)$$

Notice that $\lim_{c \rightarrow \frac{v_H + v_L}{2}} q_{sym} = \frac{1}{2}$; hence, q_{sym} is a continuous function of c on (v_L, v_H) .

Theorem 2.

- (1) q_{sym} , defined in (27), is the unique value of q such that (25) holds. $q \leq q_{sym}$ respectively implies $a_H^* \leq 1 - a_L^*$.
- (2) q_{iso} , defined in (28), is the unique value of q such that (26) holds. $q \leq q_{iso}$ respectively implies $A^* \leq q$.
- (3) $q_{iso} = q_{sym} = \frac{1}{2}$ holds if $c = \frac{v_H + v_L}{2}$. Otherwise, $c \leq \frac{v_H + v_L}{2}$ respectively implies $q_{iso} \leq q_{sym}$.

Figure 3. A graphical summary of the results in Theorem 2. Given the values of the remaining parameters, values of c and q in the region between the thinner curves are such that the seller does provide information. Values of c and q on the thicker curve correspond to information systems such that the probability of a correct signal is identical across buyer types, i.e. (25) holds; $a_H^* > 1 - a_L^*$ and $a_H^* < 1 - a_L^*$ respectively hold above and below the curve. Values of c and q on the border between the white and the gray area correspond to information systems such that the probability of trade is equal to the probability of a high-valuation match, i.e. (26) holds; $A^* > q$ and $A^* < q$ respectively hold in the gray and in the white region.

The results in Theorem 2 are illustrated in Figure 3, where the pairs $(c, q) \in (v_L, v_H) \times (0, 1)$ are classified taking (25) and (26) as the benchmarks. The values of the parameters defining the set of examples considered are $k = \frac{1}{4}$, $v_H = 1$, and $v_L = 0$. If $c < \frac{v_H + v_L}{2}$, both (25) and (26) are violated, respectively with $a_H^* > 1 - a_L^*$ and with $A^* > q$, for intervals of values of q that are greater than those for which the opposite inequalities hold; the converse is true if $c > \frac{v_H + v_L}{2}$. Additional numerical investigations, available upon request, reveal that this tendency becomes more evident as k decreases; decreases in k are also typically accompanied by smaller gaps between A^* and q , in either direction. Essentially, lower values of k are optimally exploited by choosing conditional distributions of the signal that are “farther and farther apart” between them. The changes in the relative sizes of the intervals reflect adjustments to the unconditional probabilities of the signals that reduce the cost consequences of the increased precision.

Theorem 2 can provide a valuable benchmark for the analysis of settings in which the buyer cannot observe the information system, and the seller’s incentive to provide accurate information originates either from a conditional payment to/from an outside agency, or from a reputational concern. In a setting of this type, the features of the optimal information system studied in the paper could perhaps be tempered by a relative ease with which a mis-matched product traded and/or asymmetric recommendation errors can be observed and penalized. The investigation of these trade-offs may be an interesting topic for future research.

5. Concluding remarks

I have investigated the information provision problem of a monopolist, who faces a buyer with an uncertain valuation for her product. The monopolist chooses an information system, which sends a private signal to the buyer; the cost of the information system is proportional to the mutual information between the distributions of the signal and the buyer's valuation. The equilibrium information system is noisy; in general, the probability of a correct signal differs across buyer types, and the probability of trade differs from the probability of the buyer type which makes the exchange surplus-creating. Thus, bias can be a feature of optimal information provision, even if deception is ruled out. A further, related result is that the value of information is a concave function of the quantity of information provided.

Non-neutrality reflects the idea that the agents can be expected to be relatively more careful to avoid errors that are more costly. From this point of view, the paper is related to a number of papers focusing on non-neutral, optimal information systems - such as [Burke \(2008\)](#) and [Suen \(2004\)](#). The reference to the quantity of information, measured by mutual information, makes it possible to assign a cost to general information systems, and to thereby avoid restrictions which may otherwise be necessary, for the sake of tractability and/or of comparability, such as the assumption of symmetric error probabilities. The related possibility to avoid the [Radner and Stiglitz \(1984\)](#) non-concavity problem is an additional advantage of this approach, also in view of possible future research on endogenous information structures.

The seller can effectively commit to the information system announced. Future research could investigate endogenous commitment in a model with repeated trades and/or a supervising public authority, and a record of the precision and the trustfulness of the information previously provided by the seller.

In the monopoly setting of the paper, the incentives for the seller to provide information arises from the possibility to appropriate the additional surplus created. With multiple sellers, the same incentive may result from the greater degree of product differentiation perceived by the buyers, as in [Grossman and Shapiro \(1984\)](#), [Hamilton \(2009\)](#), and [Meurer and Stahl \(1994\)](#). The relevance of the effects studied in the present paper, in the latter scenario, could also be a topic for future research.

Further topics of potential interest are the performance of different exchange mechanisms, if the precision of the information systems can be made conditional on the match-value, and the analysis of the interactions between the different parts of an information channel, if each part is subject to specific costs and/or constraints, and is managed by a distinct economic subject.

6. Appendix: Proofs

A.1. Proof of Lemma 1.

A.1.1. Part 1. $a_H^* = a_L^*$ implies $w_S^* = W$. As the buyer accepts any price $p < W$ with probability 1, if $c < W$, and if we had $p^* \in (c, W)$, the seller could certainly increase her profit $\pi(a_H, a_L, p; b_h, b_l)$, in (6), by charging a price $p \in (p^*, W)$; hence, $p^* = W$ is the only possible equilibrium price, and is actually an equilibrium price if $b_S^*(a_H^*, a_L^*, W) = 1$ also holds. To establish that this must be the case, although any $b_S^*(a_H^*, a_L^*, W) \in [0, 1]$ would in principle be compatible with (8), with $p^* = W$, notice that if we had $b_S^*(a_H^*, a_L^*, W) \in [0, 1)$, the seller could guarantee a probability of acceptance equal to 1, and thereby increase her expected profit $\pi(a_H, a_L, p; b_h, b_l)$, by charging a price p' in a suitably small left-neighborhood of p^* . The results for the case of $W < c$ are immediately obtained by noting that the seller can avoid a loss-making exchange by charging any price $p^* > W$.

A.1.2. Part 2. Notice, firstly, that $a_H^* > a_L^*$ implies $b_l^* = 0$. $b_l^* > 0$ would in fact imply $w_l^* \geq p^*$; the seller could then adopt a trivial information system with $a_H = a_L = a$, for an arbitrary $a \in [0, 1]$, and thereby force a decrease of the quantity of information provided from $I(a_H^*, a_L^*) > 0$ to $I(a, a) = 0$. At the same time, $w_l^* \geq p^*$ would imply $w_h = w_l > p^*$, under the information system (a, a) . The corresponding expected profit would then certainly be greater than that corresponding to (b_h^*, b_l^*) and p^* , which could therefore not be maximizers of $\pi(a_H, a_L, p; b_h, b_l)$, in (6).

$a_H^* > a_L^*$ also implies both $b_h^* > 0$, and $w_h^* \geq p^* > c$; otherwise, the seller could again set $a_H = a_L = a$, for an arbitrary $a \in [0, 1]$, force $I(a, a) = 0$, and thereby increase her expected profit - to 0, in this case. $p^* = w_h^*$ and $b_h^* = 1$ are then readily established by using an argument similar to that used to show that $p^* = W$ and $b_S^* = 1$, in A.1.1.

A.2. Proof of Lemma 2. The expected profit *gross* of the information cost, $qa_H(v_H - c) + (1 - q)a_L(v_L - c)$, is an affine function of (a_H, a_L) . Hence, establishing the claim is equivalent to establishing convexity of $kI(a_H, a_L)$, and therefore of $I(a_H, a_L)$. Convexity of $I(a_H, a_L)$ for general, discrete communication channels is in turn established in the second part of Theorem 2.7.4 in Cover and Thomas (1991), to which the reader is referred. To justify (15), notice that the log-sum inequality, stated as (2.100) in Cover and Thomas (1991, p. 29), holds as a *strict* inequality if at least one of the information systems considered is non-trivial. Hence, in cases of this type, the Kullback-Leibler distance between the joint distributions of the buyer type and the signal issued by the information system is *strictly* convex - see Cover and Thomas (1991, Theorem 2.7.2, p. 30).

A.3. Proof of Lemma 3. As we focus on the cases in which $a_H \geq a_L$, we must rule out optimality of the information systems featuring either $a_H = 1$ and $a_L \in [0, 1)$, or $a_H \in (0, 1)$ and $a_L = 0$. In all such cases, changes in the value of the variable on the boundary lead to a decrease both in the profit gross of the information cost, and in the quantity of information provided, which have opposite effects on $\Pi(a_H, a_L)$. The partial derivatives

$$\frac{\partial \Pi(a_H, a_L)}{\partial a_H} = q \left(v_H - c - k \log \left(\frac{a_H}{1 - a_H} \frac{1 - qa_H - (1 - q)a_L}{qa_H + (1 - q)a_L} \right) \right), \quad (\text{A.1})$$

$$\frac{\partial \Pi(a_H, a_L)}{\partial a_L} = (1 - q) \left(v_L - c - k \log \left(\frac{a_L}{1 - a_L} \frac{1 - qa_H - (1 - q)a_L}{qa_H + (1 - q)a_L} \right) \right), \quad (\text{A.2})$$

are not defined on the boundary if $a_H \in \{0, 1\}$ and if $a_L \in \{0, 1\}$, respectively. However, as $\Pi(a_H, a_L)$ is continuous over $[0, 1]^2$, and

$$\begin{aligned} \lim_{a_H \rightarrow 1} \left\{ \frac{\partial \Pi(a_H, a_L)}{\partial a_H} \Big|_{a_L \in (0, 1)} \right\} &= -\infty, \\ \lim_{a_L \rightarrow 0} \left\{ \frac{\partial \Pi(a_H, a_L)}{\partial a_L} \Big|_{a_H \in (0, 1)} \right\} &= \infty, \end{aligned}$$

the Mean Value Theorem can be invoked to conclude that both $\Pi(a'_H, a_L) - \Pi(1, a_L)$ and $\Pi(a_H, a'_L) - \Pi(a_H, 0)$ can be made positive by choosing values of a'_H and a'_L suitably close to 1 and to 0, and to thereby establish the result.

A.4. Proof of Lemma 4. The claim is established by using a standard argument - see for example the proof of uniqueness of the solution of the specular problem of minimizing a convex function on a convex set in Mangasarian (1994, Theorem 5.2.2, p. 73).

A.5. Proof of Lemma 5. It is readily verified that the signs of the derivatives $\frac{\partial q}{\partial k}$ and $\frac{\partial \bar{q}}{\partial k}$ respectively coincide with the signs of

$$\underline{g} = v_H e^{\frac{v_H}{k}} - v_L e^{\frac{v_L}{k}} - e^{\frac{v_H+v_L-c}{k}} (v_H - v_L) - c \left(e^{\frac{v_H}{k}} - e^{\frac{v_L}{k}} \right)$$

and of

$$\bar{g} = v_L e^{\frac{v_H}{k}} - v_H e^{\frac{v_L}{k}} - c \left(e^{\frac{v_H}{k}} - e^{\frac{v_L}{k}} \right) + e^{\frac{c}{k}} (v_H - v_L),$$

and are therefore both equal to 0 if $c = v_H$ and if $c = v_L$. Furthermore, convexity of the exponential function implies

$$\begin{aligned} \underline{g}|_{c=\frac{v_H+v_L}{2}} &= \frac{v_H - v_L}{2} \left(e^{\frac{v_H+3v_L}{2k}} - 2e^{\frac{v_H+v_L}{k}} + e^{\frac{3v_H+v_L}{2k}} \right) > 0, \\ \bar{g}|_{c=\frac{v_H+v_L}{2}} &= -\frac{v_H - v_L}{2} \left(e^{\frac{v_H}{k}} - 2e^{\frac{v_H+v_L}{2k}} + e^{\frac{v_L}{k}} \right) < 0. \end{aligned}$$

As both $\frac{\partial \underline{g}}{\partial c} = 0$ and $\frac{\partial \bar{g}}{\partial c} = 0$ hold at a unique (generally distinct) value of c , necessarily falling in the interval (v_L, v_H) , $\frac{\partial q}{\partial k} > 0$ and $\frac{\partial \bar{q}}{\partial k} < 0$ must hold for any $c \in (v_L, v_H)$, and therefore for any admissible value of c .

The limits of $\bar{q}$ and $\underline{q}$ as $k \rightarrow 0$ and $k \rightarrow \infty$ are readily obtained from (19) and (20), by direct calculation.

A.6. Proof of Theorem 1.

A.6.1. Part 1. By using (2), we can rearrange the first order conditions for an interior solution to (12), obtained by setting the derivatives in (A.1) and in (A.2) equal to 0, to obtain

$$\alpha_H = \frac{1}{1 + e^{-\frac{v_H-c}{k}} (A^{-1} - 1)}, \quad (\text{A.3})$$

$$\alpha_L = \frac{1}{1 + e^{-\frac{v_L-c}{k}} (A^{-1} - 1)}. \quad (\text{A.4})$$

(2) can then be rewritten as

$$A = \frac{q}{1 + e^{-\frac{v_H-c}{k}} \left(\frac{1}{A} - 1 \right)} + \frac{1-q}{1 + e^{-\frac{v_L-c}{k}} \left(\frac{1}{A} - 1 \right)}. \quad (\text{A.5})$$

The RHS of (A.5) is certainly defined for any $A \in (0, 1)$. In general, (A.5) is solved by $A = 1$ and by $A = \hat{A}$, where

$$\hat{A} = \frac{q e^{\frac{v_H-c}{k}} + (1-q) e^{\frac{v_L-c}{k}} - 1}{\left(e^{\frac{v_H-c}{k}} - 1 \right) \left(1 - e^{\frac{v_L-c}{k}} \right)}.$$

α_H and α_L in (22) and (23), obtained by setting $A = \hat{A}$ in (A.3) and (A.4), and $\hat{A}$ take values in the interval $(0, 1)$ - and are therefore compatible with an interior solution to (13), if $q \in (\underline{q}, \bar{q})$ - for $\bar{q}$ and $\underline{q}$ defined in (19) and in (20). In this case, by Lemma 4, we can conclude that the values of the respective variables are the unique solution to (13), and identify the unique equilibrium of the model. Notice that $q = q_\chi \in (\underline{q}, \bar{q})$ certainly holds, by Lemma 5, if $c = W$, for W defined in (5).

If, on the other hand, $q \notin (\underline{q}, \bar{q})$, then the first order conditions for an interior solution to (13) cannot be satisfied. Considering Lemma 3, the only possibilities left are then $(a_H^*, a_L^*) = (1, 1)$ and $(a_H^*, a_L^*) = (0, 0)$, which are both associated with $w^* = W$. The former or the latter information system is then the solution to (13), depending on whether $c < W$ or $c > W$ holds.

A.6.2. Part 2. Direct calculation reveals that if $q \in (\underline{q}, \bar{q})$, we have:

$$\frac{\partial \alpha_H}{\partial c} < 0, \frac{\partial \alpha_L}{\partial c} < 0, \frac{\partial \alpha_H}{\partial q} > 0, \frac{\partial \alpha_L}{\partial q} > 0, \frac{\partial \alpha_H}{\partial v_H} > 0, \frac{\partial \alpha_L}{\partial v_L} > 0.$$

To sign the partial derivatives

$$\begin{aligned} \frac{\partial \alpha_H}{\partial v_H} &= \frac{e^{\frac{v_H}{k}} \left(q e^{\frac{c+2v_L}{k}} - e^{\frac{c+2v_L}{k}} + e^{\frac{2v_H+v_L}{k}} - 2q e^{\frac{c+v_H+v_L}{k}} + e^{\frac{2c+v_L}{k}} + q e^{\frac{2v_H+c}{k}} - e^{\frac{2v_H+c}{k}} \right)}{k q \left(e^{\frac{v_H}{k}} - e^{\frac{c}{k}} \right)^2 \left(e^{\frac{v_H}{k}} - e^{\frac{v_L}{k}} \right)^2}, \\ \frac{\partial \alpha_L}{\partial v_L} &= - \frac{e^{\frac{v_L}{k}} \left(e^{\frac{v_H+2v_L}{k}} - q e^{\frac{c+2v_L}{k}} + 2q e^{\frac{c+v_H+v_L}{k}} - 2e^{\frac{c+v_H+v_L}{k}} - q e^{\frac{c+2v_H}{k}} + e^{\frac{2c+v_H}{k}} \right)}{k(1-q) \left(e^{\frac{c}{k}} - e^{\frac{v_L}{k}} \right)^2 \left(e^{\frac{v_H}{k}} - e^{\frac{v_L}{k}} \right)^2}, \end{aligned}$$

notice that there are unique values of q such that $\frac{\partial \alpha_H}{\partial v_H} = 0$ and $\frac{\partial \alpha_L}{\partial v_L} = 0$. Furthermore, direct calculation reveals that we have

$$\left. \frac{\partial \alpha_H}{\partial v_H} \right|_{q=\underline{q}} > 0, \left. \frac{\partial \alpha_H}{\partial v_H} \right|_{q=\bar{q}} > 0, \left. \frac{\partial \alpha_L}{\partial v_L} \right|_{q=\underline{q}} > 0, \left. \frac{\partial \alpha_L}{\partial v_L} \right|_{q=\bar{q}} > 0.$$

Hence, the “critical” values of q must both be outside $(\underline{q}, \bar{q})$, and both $\frac{\partial \alpha_H}{\partial v_H} > 0$ and $\frac{\partial \alpha_L}{\partial v_L} > 0$ necessarily hold over $(\underline{q}, \bar{q})$.

Furthermore, let (a_H^k, a_L^k) denote the values of (a_H, a_L) defined by (24) - and in particular by (22) and (23), as $k \rightarrow 0$ - for any value of $k \in \mathbb{R}_+$, given the values of the remaining parameters. If $k = 0$, we have $(a_H^0, a_L^0) = (1, 0)$. $\lim_{k \rightarrow 0} \Pi(1, 0) = qv_H$ implies $\lim_{k \rightarrow 0} \Pi(a_H^k, a_L^k) = qv_H$, and therefore $\lim_{k \rightarrow 0} (a_H^k, a_L^k) = (1, 0)$.

The results on the limits as k tends to infinity follow from the results on the behavior of the boundary values of q , in Lemma 5, and from direct calculation, in the case of $q = q_\chi$.

A.7. Proof of Theorem 2.

A.7.1. Part 1. The solutions to (25) are q_{sym} defined in (27) and, if $c \neq \frac{v_H+v_L}{2}$,

$$q' = -e^{\frac{v_H}{2k}} \left(e^{\frac{c}{k}} - e^{\frac{v_L}{k}} \right) \frac{e^{\frac{v_L}{2k}} \left(e^{\frac{v_H}{k}} - e^{\frac{c}{k}} \right) + e^{\frac{v_H}{2k}} \left(e^{\frac{c}{k}} - e^{\frac{v_L}{k}} \right)}{\left(e^{\frac{v_H}{k}} - e^{\frac{v_L}{k}} \right) \left(e^{\frac{v_H+v_L}{k}} - e^{\frac{2c}{k}} \right)}. \quad (\text{A.6})$$

The RHS of (A.6) admits no finite limit if $c \downarrow \frac{v_H+v_L}{2}$, nor if $c \uparrow \frac{v_H+v_L}{2}$. Furthermore, inspection of (A.6) immediately reveals that $q' < 0$ holds for $c \in (v_L, \frac{v_H+v_L}{2})$. This fact, together with the symmetry relation $-q'|_{c=\frac{v_H+v_L}{2}-x} = q'|_{c=\frac{v_H+v_L}{2}+x} - 1$, which holds for every $x \in (0, \frac{v_H+v_L}{2})$, implies in turn that $q' > 1$ holds if $c \in (\frac{v_H+v_L}{2}, v_H)$. Hence, $q' \notin (0, 1)$ holds for any admissible value of c , and q' is irrelevant, given our present purposes.

On the other hand, q_{sym} is always a relevant solution of (25). Given

$$(a_H^* - (1 - a_L^*))|_{q=\underline{q}} = -1, \quad (\text{A.7a})$$

$$(a_H^* - (1 - a_L^*))|_{q=\bar{q}} = 1, \quad (\text{A.7b})$$

continuity of $a_H^* - (1 - a_L^*)$ in q implies in fact the existence of some $q \in (0, 1)$ for which $a_H^* - (1 - a_L^*) = 0$; this value must then necessarily coincide with q_{sym} .

A.7.2. Part 2. Direct calculation reveals that q_{iso} in (28) is the unique value of q that solves (26), given the values of the remaining parameters. To characterize the regions defined by q_{iso} , notice that $q_{iso}|_{c=v_L} = 0$, $q_{iso}|_{c=v_H} = 1$, and $v_L < c < v_H$ guarantee

$$\frac{\partial q_{iso}}{\partial c} = -\frac{e^{\frac{v_L+c}{k}} \left(e^{\frac{v_H+v_L}{k}} - 2e^{\frac{v_H+c}{k}} + e^{\frac{2c}{k}} \right)}{k \left(e^{\frac{v_H+v_L}{k}} - 2e^{\frac{v_L+c}{k}} + e^{\frac{2c}{k}} \right)^2} > 0,$$

and therefore $q_{iso} \in (0, 1)$, for every $c \in (v_L, v_H)$. The signs of $A^* - q$ in the regions defined by q_{iso} are readily identified by using

$$\frac{\partial (A^* - q)}{\partial q} = \frac{e^{\frac{v_H}{k}} - 2e^{\frac{c}{k}} + e^{\frac{2c}{k}}}{\left(e^{\frac{c}{k}} - 1 \right) \left(e^{\frac{v_H}{k}} - e^{\frac{c}{k}} \right)} > 0,$$

which is guaranteed by $c \in (0, v_H)$.

A.7.3. Part 3. Notice that

$$(a_H^* - (1 - a_L^*))|_{q=q_{iso}} = \frac{e^{-\frac{c}{k}} \left(e^{\frac{2c}{k}} - e^{\frac{v_H+v_L}{k}} \right)}{e^{\frac{v_H}{k}} - e^{\frac{v_L}{k}}}$$

is equal to 0 if $c = \frac{v_H+v_L}{2}$, and is respectively negative and positive if $c < \frac{v_H+v_L}{2}$ and if $c > \frac{v_H+v_L}{2}$. The claim is then readily established by using (A.7), in the proof of Part 1, to identify the sub-intervals of $(0, 1)$ that must include q_{iso} , in the two cases.

References

- Anderson SP, Renault R (2006) Advertising content. *Amer Econ Rev* 96: 93-113
- Bar-Isaac H, Caruana G, Cuñat V (2010) Information gathering and marketing. *J Econ Manage Strategy* 19: 375-401
- Board O (2009) Competition and disclosure. *J Ind Econ* 57: 197-213
- Board S (2009) Revealing information in auctions: The allocation effect. *Econ Theory* 38: 125-135
- Burke J (2008) Primetime spin: Media bias and belief confirming information. *J Econ Manage Strategy* 17: 633-665
- Cabrales A, Gossner O, Serrano R (2012) Entropy and the value of information for investors. *Amer Econ Rev*, forthcoming
- Chade H, Schlee EE (2002) Another look at the Radner-Stiglitz nonconcavity in the value of information. *J Econ Theory* 107: 421-452
- Cover TM, Thomas JA (1991) *Elements of information theory*. New York: John Wiley & Sons
- Diamond PA (1971) A model of price adjustment. *J Econ Theory* 3: 156-168
- Dranove D, Jin GZ (2010) Quality disclosure and certification: Theory and practice. *J Econ Lit* 48: 935-963
- Fudenberg D, Tirole J (1991) Perfect Bayesian equilibrium and sequential equilibrium. *J Econ Theory* 53: 236-260
- Ganuza JJ (2004) Ignorance promotes competition: An auction model with endogenous private valuations. *RAND J Econ* 35: 583-598
- Ganuza J, Penalva J (2010) Signal orderings based on dispersion and the supply of private information in auctions. *Econometrica* 78: 1007-1030
- Grossman S (1981) The informational role of warranties and private disclosure about product quality. *J Law Econ* 24: 461-483
- Grossman SJ, Hart OD (1980) Disclosure laws and takeover bids. *J Finance* 35: 323-334
- Grossman GM, Shapiro C (1984) Informative advertising with differentiated products. *Rev Econ Stud* 51: 63-81

- Hamilton SF (2009) Informative advertising in differentiated oligopoly markets. *Int J Ind Organ* 27: 60-69
- Johnson JP, Myatt DP (2006) On the simple economics of advertising, marketing, and product design. *Amer Econ Rev* 96: 756-784
- Jovanovic B (1982) Truthful disclosure of information. *Bell J Econ* 13: 36-44
- Koessler F, Renault R (2012) When does a firm disclose product information? Working paper dated February 15, 2012, retrieved on March 4, 2012 at <http://www3.u-cergy.fr/renault/StrategicDisclosure.pdf>
- Levin D, Peck J, Ye L (2009) Quality disclosure and competition. *J Ind Econ* 57: 167-196
- Lewis TR, Sappington DEM (1994) Supplying information to facilitate price discrimination. *Int Econ Rev* 35: 309-327
- MacKay DJC (2003) Information theory, inference, and learning algorithms. Cambridge University Press
- Mangasarian OL (1994) Nonlinear programming. Society for Industrial and Applied Mathematics, Philadelphia
- Meurer M, Stahl DO (1994) Informative advertising and product match. *Int J Ind Organ* 12: 1-19
- Milgrom PR (1981) Good news and bad news: Representation theorems and applications. *Bell J Econ* 12: 380-391
- Milgrom P, Roberts J (1986) Relying on the information of interested parties. *RAND J Econ* 17: 18-32
- Radner R, Stiglitz J (1984) A nonconcavity in the value of information. In: Boyer M, Kihlstrom R (eds) *Bayesian models of economic theory*, Elsevier, Amsterdam, pp 33-52
- Saak AE (2006) The optimal private information in single unit monopoly. *Econ Letters* 91, 267-272
- Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27: 379-423, 623-656
- Sims CA (2003) Implications of rational inattention. *J Monet Econ* 50, 665-690
- Suen W (2004) The self-perpetuation of biased beliefs. *Econ J* 114: 377-396
- Sun M (2011) Disclosing multiple product attributes. *J Econ Manage Strategy* 20: 195-224
- Villas-Boas JM (2004) Communication strategies and product line design. *Marketing Sci* 23: 304-316

Jönköping International Business School, SE-551 11 Jönköping, Sweden
Email address: Agostino.Manduchi@ju.se