



## Pre-sale information <sup>☆</sup>

Florian Hoffmann <sup>a,\*</sup>, Roman Inderst <sup>a,b</sup>

<sup>a</sup> *Johann Wolfgang Goethe University Frankfurt (IMFS), Grüneburgplatz 1, 60323 Frankfurt am Main, Germany*

<sup>b</sup> *Imperial College London, United Kingdom*

Received 14 June 2010; final version received 8 April 2011; accepted 27 July 2011

Available online 14 October 2011

---

### Abstract

In markets as diverse as that for specialized industrial equipment or that for retail financial services, sellers or intermediaries may earn profits both from the sale of products and from the provision of pre-sale consultation services. We study how a seller optimally chooses the costly quality of pre-sale information, next to the price of information and the product price, and obtain clear-cut predictions on when information is over- and when it is underprovided, even though we find that information quality does not satisfy a standard single-crossing property. Buyers who are a priori more optimistic about their valuation end up paying a higher margin for information but a lower margin for the product when they subsequently exercise their option to purchase at a pre-specified price.

© 2011 Elsevier Inc. All rights reserved.

*JEL classification:* D42; D82

*Keywords:* Price discrimination; Single crossing; Information

---

### 1. Introduction

This paper analyzes a seller's optimal provision of information, that allows buyers to better ascertain their individual valuation, in a setting where the seller can charge separately for his product and for information. We explore how the seller optimally uses prices as well as the quality of information to better discriminate between customers with different ex-ante valuation.

---

<sup>☆</sup> We thank an associate editor, an anonymous referee, and seminar participants in Bonn, Frankfurt, and Toulouse for helpful comments, as well as Marco Ottaviani for his detailed suggestions on earlier drafts of the paper.

\* Corresponding author.

*E-mail addresses:* [fhoffmann@finance.uni-frankfurt.de](mailto:fhoffmann@finance.uni-frankfurt.de) (F. Hoffmann), [inderst@finance.uni-frankfurt.de](mailto:inderst@finance.uni-frankfurt.de) (R. Inderst).

In our baseline scenario, the seller is the monopolistic provider of both the respective product and information. Buyers differ with respect to their privately observed ex-ante valuation. Under the seller's optimal mechanism, the provision of information strictly increases in the buyer's ex-ante valuation when this is still sufficiently low, while for higher values it strictly decreases. Compared to the efficient level of information, which depends on the endogenously set product price, the seller overprovides information to high-valuation customers and underprovides information to low-valuation customers. Formally, this is the case as we show that quality of information satisfies a single-crossing property only piecewise and thus not globally for all customer types. Nevertheless, as we show, the solution to the "relaxed" problem that only considers local incentive compatibility will still be globally incentive compatible.

Our paper is linked to a recent literature that, building on Lewis and Sappington [14], analyzes the incentives of a seller or auctioneer to provide customers with information that enables them to ascertain better their true idiosyncratic preferences (cf. Johnson and Myatt [12], Esö and Szentes [6,7], Ganuza and Penalva [8], Saak [19], or Bar-Isaac et al. [3]).<sup>1</sup> While some customers may be thereby encouraged to purchase, the additional information may discourage others. What is new in our analysis is that we derive the optimal differential provision of information as a means to price discriminate.

In our model, the seller can make profits both from a margin on the sale of information and from a margin on the sale of the product. This follows Esö and Szentes [6,7], albeit by assuming that information quality is binary ("yes or no") and that costs of information provision are sufficiently small, in their models all participating types receive the same information quality. Instead, with a continuous quality of information and sufficiently high marginal costs of information provision, we can explore the role of information provision for price discrimination. By representing quality of information through a rotation of the expected value, as in Johnson and Myatt [12],<sup>2</sup> we obtain clear-cut predictions on the equilibrium provision of information and, in particular, on when there will be over- or underprovision of information in equilibrium. In fact, these results are tightly linked to the information ordering through a rotation of the expected value.

We envisage potential customers who have a privately known ex-ante valuation for a product or service. Before an order is made, but after a pre-sale contract has been concluded, the seller can provide additional information. This seems applicable to a number of settings. The provision of pre-sale services is an important part of project management, especially in industries where products or services are heavily customizable and the requirements of different customers are particularly diverse. Examples may be found in the software industry, IT services, or the market for specialized industrial equipment. Firms in these industries could either charge separately for their (pre-purchase) consulting services or only indirectly through a higher price for a subsequent order. In a procurement setting, an offer may specify both a price for a more or less detailed pre-purchase phase – e.g., carrying out a feasibility study or building a prototype – and a price for the subsequent production and delivery.<sup>3</sup>

<sup>1</sup> Less closely related are Lewis and Sappington [15] and, more recently, Krämer and Strausz [13], which focus on a moral hazard ("effort") problem in the provision of information.

<sup>2</sup> Cf. also the concept of a single-crossing ordering in Ganuza and Penalva [8].

<sup>3</sup> Such prototyping or planning plays a role in a wide range of businesses. For instance, prototyping is very common in the software industry, while an architect might offer extensive drawings or even a model of the building before starting construction work, and these services could be charged for by an hourly rate (cf. on this and further examples Terwiesch and Loch [20]). Further, firms frequently order more-or-less fully developed components – or, likewise, a variety of them

As in our model the optimal mechanism can be implemented by delegating the ultimate purchasing decision to the respective customer, by paying for the initial study or prototype the customer thus ends up acquiring an option to future delivery at a pre-specified price. From this perspective, our analysis also speaks, more generally, to the optimal provision of pre-contractual information: Though the customer pays up-front for the provision of information, he is then still free to buy at a pre-arranged price. Furthermore, with respect to business-to-customer applications, financial advisors may charge for consultation and they may also earn a margin on a subsequent transaction.<sup>4</sup> Their “menus” – as their offers are explicitly referred to by the UK’s Financial Services Authority – allow investors to (self-)select how they pay both for consultation and for investment products.<sup>5</sup> Investors can purchase different levels of consultation services, say in terms of hours of consultation or of the seniority of the respective advisor. We comment below in more detail on the role of firms’ commitment to their price structure, which in light of customer private information is essential for adopting our mechanism-design approach.

Our analysis predicts that the seller earns a higher margin on the sale of the product but a lower margin on the sale of information from buyers with a lower ex-ante valuation, compared to what he earns with buyers who have a higher valuation. Buyers with high valuation end up paying too much for information quality that is too high relative to what would be efficient, though they will then benefit from a lower purchasing price. Instead, lower-valuation buyers will pay a smaller margin on information, though what they pay for will be less than what would be efficient, and they will have to pay a higher price on a subsequent purchase.

We also investigate how these predictions change when the seller is no longer the monopolistic provider of the product, as buyers could purchase it elsewhere. One of our findings is that while a more attractive outside option reduces the price for all customers, it has a non-monotonic impact on the quality of information that they obtain in equilibrium. This highlights once more the non-standard properties of information quality to price discriminate. As noted above, in our model the quality of information satisfies a single-crossing property only piecewise and thus not globally for all customer types. Nevertheless, as we show, the solution to the “relaxed” problem that only considers local incentive compatibility will still be globally incentive compatible. What is key for this to hold is that those customer types, in terms of their prior valuation, that receive the same quality of information also have the same marginal value of information at the offered quality. Incidentally, this is also the additional global necessary incentive compatibility condition identified in Araujo and Moreira [1] for a general problem without the single-crossing property and only a single sorting variable. In our model, the solution to the “relaxed” problem satisfies their condition and, hence, can be shown to be globally incentive compatible.

The rest of the paper is organized as follows. Section 2 introduces the baseline model. The optimal menu of a monopolistic seller is characterized and discussed in Section 3. Section 4

---

with different specifications – for test runs, before going into mass production and then signing a purchase order for a large quantity.

<sup>4</sup> As independent financial advisors (IFAs) do not take ownership of the financial products, these margins are frequently paid through distribution fees or commissions, which in turn may be funded by fees that are collected from the respective investment vehicles (“loads”). Still, it is a common practice (and in some countries a legal obligation) that such distribution fees are passed on to clients when customers, instead, pay directly for advice.

<sup>5</sup> For a further description of common fee structures in the UK’s retail financial service industry see Oxera [18]. Clearly, also in other countries financial service providers frequently offer customers the choice between paying more for advice and subsequently less for products, paying nothing for advice and high product prices, or receiving no advice at all (e.g., at the time of writing such offers were available at various German online brokerages).

introduces a valuable outside option for buyers. Section 5 concludes. All omitted proofs are contained in Appendix A. Appendix B contains some additional technical results.

## 2. The baseline model

### 2.1. Information structure

To be specific, we consider a setting where a buyer can purchase a good from a seller. The buyer's expected valuation depends on how well the good's characteristics match his personal preferences. To make this assessment, the buyer can rely both on prior information as well as on information that the seller can provide. We then can suppose that the buyer's utility is additively separable,  $U := v + s$ , where  $v$  represents his prior information, while information that he obtains from the seller allows the buyer to learn about  $s$ , which satisfies  $E[s] = 0$ .<sup>6</sup>

Our main technical assumption is that these two components,  $v$  and  $s$ , are independent. This ensures, in particular, that the informativeness of the agent's ex-ante private information about his posterior valuation is constant in type and does further not depend on the quality of information provided by the seller.<sup>7</sup> Hence, the fact that a participating customer may receive additional (private) information of a certain quality from the seller does not change the degree of asymmetric information at the outset of the game. As will be shown below, we can, thus, provide a clean characterization of the optimal incentive compatible mechanism using only local incentive compatibility constraints. Instead, the most general information structure to consider would be one where  $U := v + z$ , with  $z$  arbitrarily correlated with  $v$ .<sup>8</sup> In this case, it is not surprising that one needs some restrictions on the joint distribution of the true valuation and the two signals in order to fully characterize the optimal incentive compatible mechanism using local conditions only, as already noted in the simpler settings of Esö and Szentes [6] or Courty and Li [5]. Thus, as the general setting is significantly more complex without adding important new insights, we have chosen to restrict attention to the intuitive specification with independence.<sup>9</sup>

We capture the precision of the seller's information in the following way. Based on the seller's information, the buyer forms a posterior belief about  $s$ , which we denote by  $\tilde{s}$ . From an ex-ante perspective,  $\tilde{s}$  is distributed according to  $H(\tilde{s} | \rho)$ , where the real value  $0 \leq \rho \leq \bar{\rho}$  provides a precision ordering in the sense of a mean-preserving "rotation" (cf. Johnson and Myatt [12]): For all  $\rho$  and with  $\tilde{s} \in (-\infty, \infty)$ , we suppose that  $H(\tilde{s} | \rho)$  is everywhere continuously differentiable in  $\rho$ , satisfying<sup>10</sup>

<sup>6</sup> Note that the additive structure is not an assumption, as the part of  $U$  that is unknown to the buyer ex-ante (the "error term") can always be defined as  $s := U - v$ .

<sup>7</sup> Cf. Baron and Besanko [2] or Courty and Li [5] for the concept of "informativeness" in sequential screening models.

<sup>8</sup> Following the approach in Esö and Szentes [6], making use of the probability integral transformation, one can then show that there is an equivalent formulation of the information structure where the signal disclosed by the seller is orthogonal to the buyer's ex-ante information, and still allows the buyer to execute the same Bayesian updating to compute his posterior valuation.

<sup>9</sup> Derivations for the general case, including a fully solved tractable example are given in Hoffmann and Inderst [10].

<sup>10</sup> A commonly used example that satisfies (1) is that where the seller's information generates some signal  $z$  that is "true", i.e., equal to  $s$ , with probability  $\rho$  and "fully noisy", i.e., drawn from the same distribution as  $s$ , say  $G(s)$ , with probability  $1 - \rho$ . (Note that then  $\bar{\rho} = 1$ .) When it is not possible to distinguish between the two cases, the distribution of the posterior  $\tilde{s}$  satisfies condition (1). Our working paper version, Hoffmann and Inderst [9], solves explicitly for this specification. Moreover, we show there (cf. Appendix C) that all qualitative results hold even when it is observable to

$$\frac{\partial H(\tilde{s} | \rho)}{\partial \rho} > 0 \quad \text{for } \tilde{s} < 0, \quad \frac{\partial H(\tilde{s} | \rho)}{\partial \rho} = 0 \quad \text{for } \tilde{s} = 0,$$

and

$$\frac{\partial H(\tilde{s} | \rho)}{\partial \rho} < 0 \quad \text{for } \tilde{s} > 0. \tag{1}$$

When  $\rho = 0$ , then  $H(\tilde{s} | \rho)$  is degenerate with mass one at  $\tilde{s} = E[s] = 0$ . It is convenient to stipulate that, for  $\rho > 0$ ,  $H(\tilde{s} | \rho)$  is everywhere differentiable with density  $h(\tilde{s} | \rho) > 0$ .<sup>11</sup>

The provision of more precise information comes at non-decreasing cost  $k(\rho)$ , which is continuously differentiable, with  $k(0) = 0$  and  $k(\rho) \rightarrow \infty$  as  $\rho \rightarrow \bar{\rho}$ . While it is important that the provision of information is costly, at least from some level onwards, the last assumption is for convenience as it allows us to exclude the treatment of corner solutions. Finally, we suppose that  $v \in V := [\underline{v}, \bar{v}]$  is drawn from some distribution  $F(v)$  admitting a differentiable density  $f(v)$ , strictly positive for all  $v \in V$ , and satisfying a standard monotonicity condition on the hazard rate<sup>12</sup>:

$$\frac{d}{dv} \left[ \frac{F(v)}{f(v)} \right] \geq 0 \geq \frac{d}{dv} \left[ \frac{1 - F(v)}{f(v)} \right]. \tag{2}$$

## 2.2. Mechanism

There is an initial stage,  $\tau = 1$ , at which the seller designs a mechanism. The buyer’s decision whether to participate is taken in  $\tau = 2$ , based on his private information about  $v$ . If the mechanism prescribes to do so, the seller has to provide additional information in  $\tau = 3$ . We suppose that the information quality  $\rho$  is contractible and that the thereby generated information is again only privately observed by the buyer, though all results hold when, instead, this was verifiable (cf. Appendix B). In the final period,  $\tau = 4$ , a purchase may take place and all payments are made. Both the buyer and the seller are risk neutral, and there is no discounting of payments. The cost of producing the good is given by  $\underline{c} < c_G < \bar{v}$ .

Though we allow the seller to specify in  $\tau = 1$  a general mechanism, we can without loss of generality restrict consideration to the following simple mechanism (cf. Appendix B).<sup>13</sup> In this (deterministic) mechanism, the seller’s truthful menu specifies for all participating “types”  $v$  the quality and the price of the provided information,  $\rho(v)$  and  $i(v)$ . The decision whether to then purchase the good at a price  $p(v)$  is left to the buyer. Importantly, the seller can commit to this price structure, which is essential to adopt our mechanism-design approach. We comment on this in Section 5.

While the description of the mechanism that we chose is convenient to work with, it can also be implemented by specifying a range of product prices  $p$  together with a range of non-linear

---

the buyer whether the information was truth or noise – and even when this was a verifiable event. Further, as shown in Johnson and Myatt [12], condition (1) also applies when the signal,  $z$ , is normally distributed with a precision that is monotonically related to  $\rho$ .

<sup>11</sup> Precisely, this allows to express the single-crossing property in (3) conveniently through the marginal rates of substitution.

<sup>12</sup> Note, in particular, that this assumption ensures that for all  $\kappa \in [0, 1]$  the term  $\frac{\kappa - F(v)}{f(v)}$  is weakly decreasing in  $v$ .

<sup>13</sup> The argument in Appendix B closely follows that in Esö and Szentes [6,7] and builds on the insight that, in the present setting, the buyer (agent) can only earn information rent on his prior information (cf. also Baron and Besanko [2] or Courty and Li [5]).

schedules for the price of information,  $\psi_p(\rho)$ , while now leaving it to the buyer to choose *both* how much he will pay for the product *and* the desired quality of information,  $\rho$ , at the respective price  $\psi_p(\rho)$ . Going back to the examples from the Introduction, a procurement contract could specify the volume  $\rho$  and the price  $\psi_p(\rho)$  of initial consulting services, e.g., as required to conduct a feasibility study, next to the *option* for the buyer to then place an order at a pre-specified price of  $p$ . Further, also for the example of retail financial services, the specification of such “menus” seems not to be too far fetched. To the extent that the quality of information that the seller or advisor provides is appropriately captured by the contractible time that he spends (“man days” or hours of consultation), also the specification that  $\rho$  is contracted on ex-ante seems realistic.

### 3. Analysis

For given prior valuation  $v$ , given price  $p$ , and posterior  $\tilde{s}$ , it is optimal for the buyer to purchase when  $\tilde{s} \geq p - v$ . His expected utility, when purchasing information of quality  $\rho$  at price  $i$ , is thus given by

$$u := \int_{p-v}^{\infty} [v + \tilde{s} - p] dH(\tilde{s} | \rho) - i.$$

#### 3.1. Single-crossing property in the price of the good

From  $\frac{\partial u}{\partial i} = -1$  and

$$\frac{\partial u}{\partial p} = -[1 - H(p - v | \rho)],$$

we have for all  $\rho > 0$  the following standard single-crossing property in the two contractual variables  $i$  and  $p$ :

$$\frac{\partial}{\partial v} \left[ \frac{\partial u / \partial p}{\partial u / \partial i} \right] = h(p - v | \rho) > 0. \tag{3}$$

In words, the higher is  $v$ , the more willing is the buyer to accept a higher price for information,  $i$ , in exchange for a lower price for the good,  $p$ . This is intuitive as a buyer with a higher  $v$  has, all else equal, a higher likelihood of subsequently purchasing the good.<sup>14</sup>

#### 3.2. No single-crossing property in the quality of information

In analogy to (3), we next calculate the marginal rate of substitution between the price of information  $i$  and the quality of information  $\rho$ . From this we obtain, when  $\rho > 0$ , that

$$\frac{\partial}{\partial v} \left[ \frac{\partial u / \partial \rho}{\partial u / \partial i} \right] = \frac{\partial H(p - v | \rho)}{\partial \rho}. \tag{4}$$

<sup>14</sup> As we discuss in more detail later, for the case with  $\rho = 0$  only the total transfer  $p + i$  will be pinned down. Then,  $\frac{\partial}{\partial v} \left[ \frac{\partial u / \partial p}{\partial u / \partial i} \right] = 1$  holds when the respective buyer participates and makes a purchase without obtaining additional information,  $v > p + i$ .

From (1) this is strictly positive when  $p < v$  and strictly negative when  $p > v$ . When  $p = v$ , then without additional information the buyer is just indifferent between purchasing or not. In this case, the marginal value of information is highest for the buyer. As  $|v - p|$  increases, the marginal value of information decreases. Hence, with respect to the quality of information, a single-crossing property holds only piecewise. Moreover, given that the price  $p$  will be endogenously determined, depending also on  $v$ , the range of values for which the term in (4) is positive or negative is endogenous as well.

### 3.3. The seller’s program

A buyer of “type”  $v$  who chooses the offer that was designated for “type”  $\hat{v}$  realizes the expected utility

$$u(v, \hat{v}) := \int_{p(\hat{v})-v}^{\infty} [v + \tilde{s} - p(\hat{v})] dH(\tilde{s} \mid \rho(\hat{v})) - i(\hat{v}). \tag{5}$$

Under truthtelling, we write  $u(v) := u(v, v)$ , which is non-decreasing in  $v$ . We can thus suppose the existence of some (not necessarily interior) cutoff  $v^*$  such that only types  $v \in [v^*, \bar{v}]$  participate. Buyers with  $v < v^*$  thus purchase with zero probability.

A necessary condition for truthtelling to hold, even when  $p(v)$  and  $\rho(v)$  may not be continuous (cf. Milgrom and Segal [17, Corollary 1]), is that for all  $v > v^*$  it holds that<sup>15</sup>

$$u(v) - u(v^*) = \int_{v^*}^v \frac{\partial u(v')}{\partial v'} dv' = \int_{v^*}^v [1 - H(p(v') - v' \mid \rho(v'))] dv'. \tag{6}$$

The slope of buyers’ utility,  $u(v)$ , where differentiable, is thus given by the probability with which the respective type purchases, given the price  $p(v)$  and the quality of information  $\rho(v)$ .

The seller’s expected profit, which we denote by  $\Pi$ , is the difference between the social surplus and the client’s expected surplus, i.e.,

$$\Pi := \int_{v^*}^{\bar{v}} [\omega(v) - u(v)] f(v) dv,$$

where  $\omega(v)$  denotes the social surplus realized with type  $v$ , which is given by

$$\omega(v) := \int_{p(v)-v}^{\infty} [v + \tilde{s} - c_G] dH(\tilde{s} \mid \rho(v)) - k(\rho(v)). \tag{7}$$

Using (6) and integration by parts,  $\Pi$  can then be written as

$$\Pi = \int_{v^*}^{\bar{v}} \hat{\omega}(v) f(v) dv, \tag{8}$$

<sup>15</sup> The assumptions of Corollary 1 in Milgrom and Segal [17] are satisfied as  $V$  is compact and as, with a slight abuse of notation, the buyer’s utility  $u(v, \rho, p, i)$  is absolutely continuous and differentiable, with  $\frac{\partial u}{\partial v} \in [0, 1]$  for all possible  $\rho, p, i$ .

where  $\hat{\omega}(v)$  denotes the virtual surplus

$$\hat{\omega}(v) := \omega(v) - \left( \frac{1 - F(v)}{f(v)} \right) [1 - H(p(v) - v \mid \rho(v))]. \tag{9}$$

Ignoring the issue of global incentive compatibility for the moment and differentiating  $\hat{\omega}(v)$  in (9) with respect to  $p(v)$  for pointwise optimization, we have that independent of the choice  $\rho(v) \geq 0$  the optimal unique price of the good satisfies

$$p(v) = c_G + \frac{1 - F(v)}{f(v)}. \tag{10}$$

The mark-up above the seller’s cost,  $c_G$ , distorts the buyer’s subsequent purchasing decision. Note that it is independent of the quality of the buyer’s additional information. The intuition for this result is as follows. What matters for screening is how informative the customer’s ex-ante private information is about his true valuation; or, more precisely, how the agent’s private knowledge about the distribution of his posterior valuation changes across types. This is analogous to the notion of “informativeness” in standard sequential screening models as, e.g., Baron and Besanko [2] or Courty and Li [5]. With  $v$  and  $s$  independent, however, this informativeness measure does not depend on  $\rho$ , resulting in the expression in (10).<sup>16</sup>

Given the hazard rate condition (2),  $p(v)$  then is weakly decreasing in  $v$ . When  $p(v)$  satisfies (10), we can denote the unique buyer type satisfying  $p(v) = v$  by  $\tilde{v}$ <sup>17</sup>:

$$p(\tilde{v}) = \tilde{v}. \tag{11}$$

When some buyers are excluded, as  $v^* > \underline{v}$ , we show in Proposition 1 that it still holds that  $v^* < \tilde{v}$ , such that type  $\tilde{v}$  indeed buys with strictly positive probability. As a consequence, we have from (4) that a single-crossing property in the quality of information holds indeed only piecewise over the set of participating types, namely over the subsets  $[v^*, \tilde{v}]$  and  $(\tilde{v}, \bar{v}]$ . We comment more on this below.

### 3.4. Information quality

When  $\rho(v) > 0$ , then the first-order condition for maximizing  $\hat{\omega}(v)$  with respect to  $\rho(v)$  can be transformed to

$$\frac{\partial}{\partial \rho} \int_{p(v)-v}^{\infty} [v + \tilde{s} - p(v)] dH(\tilde{s} \mid \rho) \Big|_{\rho=\rho(v)} = k'(\rho(v)), \tag{12}$$

where  $p(v)$  is substituted from (10). Note that the left-hand side of (12), which represents the seller’s first-order condition, also equals the buyer’s marginal value of (more) information, at given price  $p(v)$ . Hence, at the seller’s optimal price  $p(v)$ , the seller’s marginal benefits of increasing information quality is equal to the marginal value of information to the buyer. As we discuss below, this observation will be key to ensure global incentive compatibility.

<sup>16</sup> However, while it is sufficient, independence of  $v$  and  $s$  is not a necessary assumption. A specification where both components are normally distributed and correlated but where the optimal screening price  $p(v)$  still is independent of information quality is given in Hoffmann and Inderst [10].

<sup>17</sup> Existence of  $\tilde{v} \in (\underline{v}, \bar{v})$  follows from  $\underline{v} < c_G < \bar{v}$ .

Note next that from the rotation in (1) and as  $v - p(v)$  is strictly increasing in  $v$ , the left-hand side of (12) is increasing in  $v$  when  $v < \tilde{v}$  and decreasing when  $v > \tilde{v}$ . Consequently, even when  $\rho(v)$  is not everywhere uniquely determined, though it is so almost everywhere, any selection of optimal values  $\rho(v)$  is hump-shaped: It is increasing for  $v < \tilde{v}$  and decreasing for  $v > \tilde{v}$  (and strictly so where positive).<sup>18</sup> From (12), a sufficient condition for  $\rho(v) > 0$  to hold at least for some  $v$  is that

$$\frac{\partial}{\partial \rho} \int_0^\infty \tilde{s} dH(\tilde{s} | \rho) \Big|_{\rho=0} > k'(0), \tag{13}$$

which is also a necessary condition when  $k''(\rho) > 0$ . In what follows, we assume that (13) holds.

We now compare the derived optimal value  $\rho(v)$  with a benchmark of efficiency. For this we take the distorted price  $p(v)$  as given and take as a benchmark the efficient provision of information conditional on  $p = p(v)$  for the respective type  $v$ . To see how the equilibrium provision of information compares with this benchmark of (constrained) efficiency, it is useful to revisit Eq. (9), which determines the virtual surplus  $\hat{\omega}(v)$ . Differentiating with respect to  $\rho(v)$  and comparing this to the derivative of the true surplus,  $\omega(v)$ , while using the definition of a rotation in (1), we have  $\partial \hat{\omega}(v)/\partial \rho(v) < \partial \omega(v)/\partial \rho(v)$  for  $v < \tilde{v}$ ,  $\partial \hat{\omega}(v)/\partial \rho(v) = \partial \omega(v)/\partial \rho(v)$  for  $v = \tilde{v}$ , and  $\partial \hat{\omega}(v)/\partial \rho(v) > \partial \omega(v)/\partial \rho(v)$  for  $v > \tilde{v}$ . This comparison of the marginal impact on the true and the virtual surplus immediately implies that there is underprovision of information when  $v < \tilde{v}$ , efficient provision when  $v = \tilde{v}$ , and overprovision when  $v > \tilde{v}$ .<sup>19</sup> Moreover, this makes also immediate that our characterization of when information is under- or overprovided is tightly linked to the use of the rotation structure in (1) as a measure of informativeness.<sup>20</sup>

Inspection of the virtual surplus function in (9) also makes transparent when there will *always* be an efficient provision of information: at both  $v = \tilde{v}$ , i.e., for a buyer with an intermediate ex-ante valuation, and  $v = \bar{v}$ , i.e., for the buyer with the highest ex-ante valuation. This observation further highlights the non-standard properties of information provision as a tool for price discrimination and rent extraction, which arise from the fact that the single-crossing property holds only piecewise (cf. (3)).

So as to extract more of the information rent of higher-valuation buyers, the seller wants to make  $u(v)$  flatter. As for lower  $v$  the marginal value of information increases in  $v$ , this is achieved through *decreasing* information quality, while for higher  $v$ , where the marginal value of information decreases in  $v$ , this is achieved through *increasing* information quality. Intuitively, for  $v < \tilde{v}$ , a marginal increase in information quality helps reducing the probability of committing the error of not purchasing the good despite  $U > p(v)$ , thereby increasing the overall probability of a purchase. Hence, in order to reduce the rents to be paid to higher types, the seller optimally *decreases* information quality. For  $v > \tilde{v}$ , a marginal increase in information quality reduces the overall probability of a purchase (by reducing the probability of an erroneous purchase despite

<sup>18</sup> Our restriction to a deterministic mechanism thus rules out only the possibility that for a measure-zero set of  $v$  there could be randomization over  $\rho(v)$  and  $i(v)$  (cf. also the general mechanism approach in Appendix B).

<sup>19</sup> Proposition 1 makes precise when under- or overprovision hold strictly.

<sup>20</sup> Incidentally, the distortions in information provision follow the same pattern also if the first-best (thus assuming efficient consumption or, likewise, that  $p(v) = c_G$ ) is used as a benchmark. In particular, there still is underinvestment for small  $v$  and overinvestment in case  $v$  is large enough.

$U < p(v)$ ), and, thus the seller optimally *increases* information quality for the purpose of rent reduction.<sup>21</sup>

**Proposition 1.** *The optimal offer of a seller who can charge discriminatory prices both for a product and for the prior provision of information is characterized as follows. There are two unique cutoff valuations  $\underline{v} \leq v^* < \tilde{v} < v^{**} \leq \bar{v}$ . Buyers with ex-ante valuation  $v > v^*$  purchase with positive probability at the respective unique price (10). Buyers  $v \in (v^*, v^{**})$  obtain from the seller additional information of quality  $\rho(v) > 0$ , which solves (12) and which is almost everywhere uniquely determined. Information quality is strictly increasing in the ex-ante valuation  $v$  when  $v^* < v < \tilde{v}$ , where it is underprovided, while it is strictly decreasing when  $\tilde{v} < v < v^{**}$ , where it is overprovided. Buyers  $v = \tilde{v}$  and  $v \in [v^{**}, \bar{v}]$  always receive the respective efficient level of information quality (though this is zero for  $v \in [v^{**}, \bar{v}]$ , when  $v^{**} < \bar{v}$ ).*

### 3.5. Margins

Interestingly, even though from Proposition 1 information provision is non-monotonic in  $v$ , we obtain a monotonic prediction for the margin that the seller earns from the sale of the product,  $p(v) - c_G$ , as well as from the sale of information,  $i(v) - k(\rho(v))$ .<sup>22</sup>

**Corollary 1.** *The seller's margin from the sale of the good is (weakly) decreasing in the buyer's ex-ante valuation  $v$ , while the margin from the sale of information is (weakly) increasing.*

When  $v^*$  is interior,  $v^* > \underline{v}$ , and thus pinned down by the requirement that the virtual surplus is zero,  $\hat{\omega}(v^*) = 0$ , then with the lowest type  $v = v^*$  the seller's margin for the provision of information is zero:  $i(v^*) = k(\rho(v^*))$ .<sup>23</sup> When also the highest type  $\bar{v}$  purchases information,  $\rho(\bar{v}) > 0$ , then this is sold at a positive margin,  $i(\rho(\bar{v})) > k(\rho(\bar{v}))$ . Taken together, as also  $p(\bar{v}) = c_G$  holds, in this case the seller earns profits from the highest-valuation buyer  $\bar{v}$  *only* through the sale of information and profits from the lowest-valuation buyer  $v^*$  *only* through the sale of the product. More generally, from Corollary 1 the respective margins change monotonically with the buyer's ex-ante valuation.

Recall at this point from the Introduction the brief discussion of charges in retail financial services. Proposition 1 and Corollary 1 show how making some buyers pay more than others (and more directly) for consultation can allow a seller to better price discriminate. This contrasts with the common perception that sellers may want to shift payments away from up-front charges,  $i(v)$ , in order to make their pricing less transparent.

### 3.6. Global incentive compatibility

The mechanism characterized in Proposition 1 proves to be globally incentive compatible even though a single-crossing property in the quality of information does not hold for all participating

<sup>21</sup> For brevity's sake, the statement of Proposition 1 does not include the characterization of  $v^*$ ,  $v^{**}$ , and of the price of information  $i(v)$ . This is relegated to the proof of Proposition 1 in Appendix A. Note also that when  $v^{**} < \bar{v}$ , a positive measure  $(v^{**}, \bar{v}]$  of types does not purchase additional information and will thus purchase the good with probability one. Clearly, in this case only the total sure payment  $i(v) + p(v)$  is pinned down uniquely.

<sup>22</sup> Breaking up the customer's possible payments in this way, i.e., as a charge for information and a price for the product, seems natural, even though, more generally,  $i(v)$  represents an entry payment to participate in the seller's mechanism.

<sup>23</sup> This holds irrespective of whether  $\rho(v^*)$  is strictly positive or zero.

types, but only piecewise. To establish this, the proof of [Proposition 1](#) makes use of the following two properties. First, there are only two regions, namely  $v < \tilde{v}$  and  $v > \tilde{v}$ , over which a single-crossing property is satisfied. Second, the seller's first-order condition (12) for  $\rho(v)$ , which is obtained from  $\partial \hat{\omega}(v)/\partial \rho = 0$ , also ensures that buyer types  $v < \tilde{v} < v'$  who receive the same quality of information,  $\rho(v) = \rho(v') = \rho$ , have the same marginal value of information at  $\rho$  (and thus also the same marginal rate of substitution between  $\rho$  and  $i$ ).

Incidentally, this is also the additional global necessary incentive compatibility condition identified in Araujo and Moreira [1] for a general problem without the single-crossing property and only a single sorting variable. In our model, the solution to the “relaxed” problem satisfies their condition and, hence, can be shown to be globally incentive compatible.

### 3.7. Discussion: Revisiting the result on over- and underprovision of information

In our setting, the seller can provide buyer-specific information in a discriminatory way through the menu. In contrast, papers such as Lewis and Sappington [14] and Johnson and Myatt [12] analyze the opposite extreme where the seller has to provide all buyers with the same quality of information and, in addition, cannot charge for it (cf. also Ganuza and Penalva [8] for an auction setting).

As noted in the Introduction, Esö and Szentes [6,7] also allow the seller to charge for information, both in an auction and in a bilateral setting of advice. Assuming that the seller's cost of providing a fixed amount of information is sufficiently small, they show that this information is provided to all participating types with probability one (the “efficient level”). Note that in these papers, where information quality is binary, the seller could commit to randomize between information revelation and no information revelation. Still, the implied linearity gives rise to an “all-or-nothing” supply of information. In contrast, our model with continuous information quality (together with sufficiently high marginal costs of information provision) generates an interior solution for information quality, so that only a buyer with some intermediary *ex-ante* valuation,  $v = \tilde{v}$ , always obtains a quality of information that is both strictly positive and efficient. Optimal price discrimination requires, instead, that there is overprovision for higher and underprovision for lower types. Let us note once more that these results continue to hold in case  $\tilde{s}$  was verifiable and could thus be explicitly included in the mechanism (cf. [Appendix B](#)). Hence, as in Esö and Szentes [6,7], the buyer can be maintained at the level of her information rent at the beginning of the game, i.e., she does not earn any (additional) information rent on the signal controlled by the seller, irrespective of the provided quality of information. By distorting information quality relative to the efficient level, the seller is able to reduce the information rent the customer earns on his *ex-ante* information.

The main novelty in our model is thus that we can explore the role of information provision for price discrimination. The preceding analysis explores this in a model of monopoly pricing. Over the last decade, the theory of price discrimination and adverse selection has provided also a unifying treatment of the case where buyers have a valuable, possibly type-dependent outside option (cf. Biglaiser and Mezzetti [4], Maggi and Rodriguez [16], and, most recently, Jullien [11]). We next extend our theory to such a framework.

## 4. Outside good

In contrast to our baseline case, we now suppose that the same good may also be purchased elsewhere at some given (market) price  $\underline{v} < m < \bar{v}$ , where we will allow both for  $m \geq c_G$  and for

$m < c_G$ . When  $m > c_G$  holds strictly, the seller enjoys, next to his monopolistic position for the provision of information, also a cost advantage for the provision of the good.

#### 4.1. The seller's problem

We can again restrict consideration to a simple mechanism that prescribes for all participating types a contract  $\{i(v), \rho(v), p(v)\}$ .<sup>24</sup> Note next that the outside option generates type-dependent reservation values equal to

$$w(v) := \max\{v - m; 0\}. \quad (14)$$

The particular shape of  $w(v)$  in (14) allows to restrict attention to a relaxed program that considers only the participation constraints at the boundaries:  $u(v^*) \geq w(v^*)$  and  $u(\bar{v}) \geq w(\bar{v})$ . This follows as from  $du/dv = 1 - H(p(v) - v | \rho(v))$  (at points of differentiability) we have for all participating types  $0 < du/dv \leq 1$ , with equality only when  $\rho(v) = 0$  and  $v > p(v)$ , while  $dw/dv$  is from (14) either zero (for  $v < m$ ) or one (for  $v > m$ ).<sup>25</sup>

As will be intuitive, under the optimal mechanism, the seller will produce the good himself in case  $c_G < m$  and delegate production to the market if  $c_G > m$ . (When  $c_G = m$ , both specifications are fully payoff equivalent). Hence, one may equivalently think of the present setting as one where the information provider searches for the most (cost-)efficient product on the market and subsequently sells it on to the customer. In terms of the applications discussed in the Introduction, this mirrors the practice of so-called “integrated service providers”. For the following analysis it will, thus, be convenient to use the notation  $c := \min\{c_G, m\}$ .

When the participation constraint is slack at  $\bar{v}$ ,  $u(\bar{v}) > w(\bar{v})$ , the seller's program is identical to that analyzed for Proposition 1.<sup>26</sup> As we show below, and as is intuitive, this case applies when the outside option is sufficiently unattractive as  $m$  is sufficiently large. At the opposite extreme, which may apply when  $m$  is sufficiently small relative to the seller's own cost  $c_G$ , the participation constraint is slack at  $v^*$ .<sup>27</sup> The third and final case is that where the constraint binds at both boundaries.

#### 4.2. Characterization

We can show (cf. the proof of Proposition 2, but also Jullien [11] or Esö and Szentes [7]) that the virtual surplus is now, more generally, given by<sup>28</sup>

$$\hat{\omega}(v) := \omega(v) - \frac{\kappa - F(v)}{f(v)} [1 - H(p(v) - v | \rho(v))], \quad (15)$$

<sup>24</sup> We only consider the case where the seller can force a participating customer to purchase from him, instead of from the market. (That is, a purchase from the market is verifiable.) We can show that relaxing this assumption changes the seller's program (only) by imposing the additional constraint that  $p(v) \leq m$ .

<sup>25</sup> Note that we choose not to apply optimal control theory, as explored with type-dependent reservation values in Jullien [11]. There, the standard conditions of optimality are obtained by restricting the control variable to the set of piecewise continuous functions. As we have seen in the baseline case, however, this may not hold for the optimal  $\rho(v)$ .

<sup>26</sup> In fact, as we show in the proof of Proposition 2, it then must also hold that  $u(v^*) = 0$ , i.e., that  $v^* < m$ , which intuitively follows from  $dw/dv = 1$  for  $v > m$  together with  $du/dv \leq 1$  (strictly when  $\rho(v) > 0$ ).

<sup>27</sup> As is easily seen, in this case, we must have  $v^* = \underline{v}$ .

<sup>28</sup> The true surplus is still given by the expression in (7), after substituting  $c$  for  $c_G$ .

where  $\kappa = 1$  corresponds to our original case, with  $u(\bar{v}) > w(\bar{v})$ ;  $\kappa = F(v^*)$  corresponds to the opposite extreme, where the participation constraint is slack at the bottom,  $u(v^*) > w(v^*)$ ; and values  $F(v^*) < \kappa < 1$  correspond to the case where both  $u(v^*) = w(v^*)$  and  $u(\bar{v}) = w(\bar{v})$ . Intuitively, to extract more information rent, from (15) the seller would want to make  $u(v)$  flatter for all  $v < v_0$ , where  $v_0$  satisfies

$$F(v_0) = \kappa, \tag{16}$$

but to make it steeper for all  $v > v_0$ .<sup>29</sup> Differentiating (15) with respect to  $p(v)$  obtains

$$p(v) = c + \frac{\kappa - F(v)}{f(v)}, \tag{17}$$

which thus strictly exceeds the seller's cost for  $v < v_0$  and strictly lies below his cost for  $v > v_0$ . Further, when  $\rho(v) > 0$ , then the first-order condition for maximizing  $\hat{\omega}(v)$  in (15) with respect to  $\rho(v)$  can be transformed to still obtain (12), however, with (17) substituted as the optimal pricing schedule.<sup>30</sup>

**Proposition 2.** *Suppose that instead of signing a contract with the seller-advisor, a buyer can initially also purchase a product from the market at price  $m$ , albeit without obtaining additional information. Then, the seller's optimal offer is still unique (albeit only almost everywhere in terms of  $\rho(v)$ ), with the product price  $p(v)$  given by (17) and all values  $\rho(v) > 0$  still solving (12).*

### 4.3. Comparative analysis

In the following, we are interested in how the attractiveness of the outside option affects the optimal mechanism, in particular the quality of information that customers purchase in equilibrium. We restrict attention to the case where the seller enjoys, next to his monopolistic position for the provision of information, also a cost advantage for the provision of the good, i.e.,  $c_G \leq m$ . Note that a change in  $m$  does not directly affect the (marginal) value of information. As  $m \geq c_G$ , in equilibrium, the outside option will not even be taken up with positive probability by any customer. A change in  $m$  will, however, affect the value of information indirectly through the effect that it has on the product price,  $p(v)$ . While the latter will be monotonic, we find that the quality of information can change non-monotonically in  $m$ .

From (17), for any  $v$ , the price is lower when also  $\kappa$  is lower, which – as we show – is the case when the outside option becomes more attractive as  $m$  decreases. The implications for the equilibrium information quality are, however, more involved. To get some intuition for the subsequent characterization, it is important to recall the following two observations. First, the seller's optimal choice of  $\rho(v)$ , which is still given by (12) after substituting for  $p(v)$  from (17), equates the buyer's marginal value for information with the seller's marginal cost. Second, the buyer's marginal value for information is strictly decreasing in  $|v - p(v)|$  and thus highest at  $\bar{v}$ , where  $\bar{v} = p(\bar{v})$ .<sup>31</sup> We now use these two observations to determine how the provision of information changes in the value of buyers' outside option.

<sup>29</sup> Note again that when  $\kappa = 1$ , which holds when  $m$  is large, then with  $v_0 = \bar{v}$  we are back to the original case, where he seller wants to make  $u(v)$  flatter for all  $v$ .

<sup>30</sup> For brevity, Proposition 2 does not include a characterization of  $v^*$ ,  $v^{**}$  and  $i(v)$ .

<sup>31</sup> While  $\bar{v}$  now clearly depends on  $\kappa$ , for ease of exposition we do not make this dependency explicit in the main text.

When a buyer has a sufficiently high ex-ante valuation  $v$ , then as  $p(v)$  decreases, given a lower  $m$  and thus lower  $\kappa$ , this always increases  $|v - p(v)|$ , as then  $v - p(v) > 0$ . By the preceding remarks this leads to lower information quality  $\rho(v)$ . The opposite happens for buyers who have sufficiently low ex-ante valuation. There, as always  $v - p(v) < 0$ , a lower  $p(v)$  always reduces  $|v - p(v)|$ . This leads to always higher information quality. Finally, for all buyers with an intermediate ex-ante valuation, as  $p(v)$  decreases, then  $|v - p(v)|$  first decreases, when  $p(v)$  is still high such that  $v - p(v) < 0$ , and then increases, when  $p(v)$  is sufficiently low such that  $v - p(v) > 0$ . For intermediate values  $v$  there is thus a *non-monotonic* relationship between the attractiveness of the substitute good (lower  $m$ ) and the quality of information that these buyers purchase in equilibrium.<sup>32</sup>

**Corollary 2.** *Under the seller's optimal offer, as characterized in Proposition 2, and when  $m \geq c_G$ , the following comparative results hold as the outside option becomes more attractive (lower  $m$ ):*

- (i) *As this leads to a decrease in  $\kappa$ , all product prices  $p(v)$  are lower and there is more participation (lower  $v^*$ ) (both strictly when  $F(v^*) < \kappa < 1$ ).*
- (ii) *There is always more information provided to buyers with sufficiently low valuation, provided that they participate, and always less information to buyers with sufficiently high valuation. For buyers with an intermediate ex-ante valuation, information quality  $\rho(v)$  first increases and then decreases as the outside option becomes more attractive.*

It should be noted that these changes in information provision that are induced by changes in  $m$  do not follow directly from a change in the value of information. Instead, as we discussed in detail above for the different cases, a change in the value of information is only induced by the concurrent change in  $p(v)$ .

#### 4.4. Efficiency of information provision

Again, we use as a benchmark the efficient provision of information conditional on the decision with prices  $p = p(v)$ . Whether, for a given type  $v$  who receives information, there is over- or underprovision, depends now on how  $v$  is located relative to *both*  $\tilde{v}$ , which is the type with the highest marginal valuation for information, *and* relative to  $v_0$ , which is the type that separates the interval  $[v^*, v_0)$ , where the seller wants to make  $u(v)$  flatter to reduce information rent, from the interval  $(v_0, \tilde{v}]$ , where the seller wants to make  $u(v)$  steeper. The resulting implications are now immediate from inspection of the virtual surplus in (15). Recall from this that the term

$$\left[ \frac{\kappa - F(v)}{f(v)} \right] [1 - H(p(v) - v | \rho(v))] \quad (18)$$

captures the difference between surplus and virtual surplus. This term is central to the overall understanding of how information provision serves as a tool to price discrimination under different circumstances, which is the key focus of this paper. To reduce this term, thereby increasing his profits, the seller wants to decrease the probability of trade, as given by the second part in (18),

<sup>32</sup> The proof of Proposition 2 in Appendix A derives explicitly the boundaries for buyers' valuation  $v$  for which the different cases in assertion (ii) apply and information quality thus changes either monotonically or non-monotonically in  $m$ .

when the first part in (18) is positive ( $v < v_0$ ) and to increase the probability of trade, when the first part is negative ( $v > v_0$ ). How the probability of trade is in turn affected by higher information quality depends on whether  $v < \tilde{v}$  (higher probability) or  $v > \tilde{v}$  (lower probability).

**Corollary 3.** *When  $\kappa$  decreases as the substitute good becomes more attractive (lower  $m$ ), the efficiency of information provision changes as follows:*

- (i) *For high  $\kappa$ , where we have  $v^* < \tilde{v} < v_0 \leq \bar{v}$  (and where  $v_0 = \bar{v}$  holds for  $\kappa = 1$ ), there is underprovision of information for low-valuation buyers  $v < \tilde{v}$  and overprovision for all  $\tilde{v} < v < v_0$ . At  $v = \tilde{v}$  and  $v = v_0$  there is efficient information provision. When  $v_0 < \bar{v}$  holds strictly, then there is also underprovision for all  $v > v_0$ .*
- (ii) *For one intermediate value of  $\kappa$ , we have  $v^* < \tilde{v} = v_0 < \bar{v}$ , such that for all  $v \neq \tilde{v} = v_0$  there is underprovision, while for  $v = \tilde{v} = v_0$  there is efficient information provision.*
- (iii) *For low  $\kappa$ , where we have  $v^* \leq v_0 < \tilde{v} < \bar{v}$  (and where  $v^* = v_0 = \underline{v}$  holds for  $\kappa = 0$ ), there is underprovision of information for high-valuation buyers  $v > \tilde{v}$  and overprovision for all  $v_0 < v < \tilde{v}$ . At  $v = \tilde{v}$  and  $v = v_0$  there is efficient information provision. When  $v_0 > v^*$  holds strictly, then there is also underprovision for all  $v < v_0$ .*

Hence, when buyers can purchase the product also elsewhere, albeit without information, a more complex picture emerges. Still, what holds true for all cases (i)–(iii) in Corollary 3 is that always some interior type, which may be  $\tilde{v}$  or  $v_0$  (unless they coincide in case (ii)), receives the efficient information quality.

Finally, it is immediate to show that Corollary 1 still holds when we introduce a valuable outside option.<sup>33</sup> Hence, provided that they receive information, the seller's respective margin  $i(v) - k(\rho(v))$  is still higher with high-valuation buyers, while his margin from a subsequent sale of the product,  $p(v) - c$ , is lower (and even negative when  $v > v_0$ , as given in (16)).

## 5. Conclusion

In our model, a seller can earn profits both by charging a positive margin on the good that he sells and by charging for the provision of information that helps buyers to better determine the good's suitability, given their specific needs. Crucially, we allow the seller to contract on the provision of information with various quality, e.g., as measured in days spent on preparing a pre-purchase (feasibility) study or in hours of face-to-face consultation. One result that robustly arises throughout the various scenarios that we consider is the following: Compared to buyers with a low ex-ante valuation, the seller makes from buyers with higher ex-ante valuation a higher margin on the sale of information, when it is provided, and a lower margin on the sale of the product. Somewhat loosely speaking, low-valuation buyers, who are less likely to ultimately purchase the product in equilibrium, thus seem to be paying for information (relatively more) "indirectly", through a higher margin for the product.

In equilibrium, information quality is hump-shaped in buyers' ex-ante valuation. An interesting implication of the fact that information provision does not satisfy a single-property over all buyer types, but only piecewise, is that there are (generically) two types (or intervals) where information provision is efficient. The particular nature of information provision as a means of

<sup>33</sup> Formally, this follows immediately as the proof of Corollary 1 only relies on monotonicity of  $p(v)$ .

price discrimination is further underlined by the observation that, as the buyers' outside option becomes more attractive, the quality of information changes non-monotonically for buyers with an intermediate ex-ante valuation, for whom the marginal value of information is thus highest.

One purpose of the present paper is to highlight the non-standard features of information as a tool for price discrimination in a model where customers differ in their ex-ante valuation of the product. This is just one particular information structure one might want to consider. Alternatively, there could be ex-ante heterogeneity with respect to customers' initial valuation uncertainty, as discussed, e.g., in Courty and Li [5]. Using their framework, our own calculations show that the techniques developed in the present paper also apply in this setting. However, the characterization is remarkably different. When customers differ in their valuation uncertainty, their marginal value of new information changes monotonically. Then, both the efficient and the equilibrium provision of information quality are monotonic. More importantly, as then a standard global single-crossing property holds, we obtain for the monopolistic seller also the standard screening result with underprovision of information quality for all but the highest type.<sup>34</sup> It seems natural and interesting to explore how the seller would optimally price discriminate when he faces customers who may differ both in their ex-ante valuation and in their valuation uncertainty, which then gives rise to a two-dimensional screening problem.

Finally, as we remarked earlier, to adopt a mechanism-design approach the seller needs to commit to the particular price structure, even when there is scope for mutually beneficial renegotiations. Such a scope for renegotiation arises at various points in the described mechanism. For instance, a customer with low ex-ante valuation may simply refrain from participating even though this would be efficient when the product was supplied at cost and information was efficiently provided. Likewise, after receiving information there may be gains from trade even though the customer refrains from a purchase given a price above cost; while when the price is below cost, there may be efficiency gains when the customer would refrain from purchasing. While there is substantial literature that analyzes dynamic pricing without commitment ("Coase Conjecture"), it would be interesting to see how the analysis unfolds when, as in the present problem, the seller controls both the price and the provision of information, which is left for future research.

## Appendix A. Omitted proofs

**Proof of Proposition 1.** Over  $v \geq v^*$ , the optimal choice of  $p(v)$  and  $\rho(v)$  for the "relaxed" program maximizes  $\hat{\omega}(v)$ . Where  $\rho(v) > 0$ , this obtains the unique and (weakly) decreasing characterization in (10). We extend the definition of  $p(v)$ , as in (10), to all  $v \geq v^*$ , irrespective of the choice of  $\rho(v)$ . Note next that  $\frac{\partial^2 \hat{\omega}}{\partial \rho \partial v}$  has the same sign as  $p(v) - v$ . Monotonicity of any optimal selection  $\rho(v)$  for  $v < \tilde{v}$  and  $v > \tilde{v}$ , as asserted in the proposition, follows then from standard monotone comparative statics results (cf. Vives [21, Theorem 2.3]).

At  $v = \tilde{v}$ , where  $p(v) = v$ , note next that for all  $\rho > 0$  the marginal value of information is constant. As  $k(\rho)$  is continuously differentiable, from (13) this strictly exceeds  $k'(\rho)$  for all values of  $\rho$  that are sufficiently close to zero, implying that  $\rho(\tilde{v}) > 0$  holds strictly. (Clearly, this characterization is only of relevance when  $\tilde{v} > v^*$ .) As  $\hat{\omega}(v) = 0$  holds for all  $v < \tilde{v}$  when  $\rho(v) = 0$  and as  $\hat{\omega}(v)$  is clearly continuous, this has the following implication: when  $\tilde{v} < v^*$ ,

<sup>34</sup> The respective calculations are provided in Hoffmann and Inderst [10].

then for a sufficiently small but strictly positive  $\varepsilon$  it (still) holds that  $\rho(v) > 0$  and  $\hat{\omega}(v) > 0$  for all  $v \in (\tilde{v} - \varepsilon, \tilde{v})$ . Next, for types  $v > \tilde{v}$ , monotonicity of  $\rho(v)$  implies existence of the asserted cutoff  $v^{**}$ . (We do not claim that  $v^{**} < \bar{v}$  holds strictly.) From the preceding remarks it follows immediately that there exists a uniquely optimal choice satisfying  $v^* < \tilde{v}$ . This is equal to the lowest value of  $v$  such that  $\rho(v) > 0$  holds for all higher values of  $v$  up to  $\tilde{v}$ . When this value of  $v$  is interior with  $v^* > \underline{v}$ , then  $\hat{\omega}(v^*) = 0$ .

To conclude the characterization, note finally that the price of information  $i(v)$  for  $v < v^{**}$  is from (6) uniquely determined by

$$i(v) = \int_{p(v)-v}^{\infty} [v + \tilde{s} - p(v)] dH(\tilde{s} | \rho(v)) - \left[ \int_{v^*}^v [1 - H(p(v') - v') | \rho(v')] dv' \right], \tag{19}$$

while incentive compatibility at  $v^{**}$  uniquely pins down the price without information, which we denote by  $q$ :

$$\lim_{v \rightarrow v^{**}} \left[ \int_{p(v)-v}^{\infty} [v + \tilde{s} - p(v)] dH(\tilde{s} | \rho(v)) - i(v) \right] = v^{**} - q,$$

where we use that the limit exist from monotonicity of  $\rho(v)$ .

*Efficiency.* To prove next the asserted comparison with the efficient choice of  $\rho(v)$ , which for given  $p(v)$  maximizes  $\omega(v)$ , define  $\Omega(\rho(v), a) := \omega(v) - a[\hat{\omega}(v) - \omega(v)]$  and observe that, evaluated at  $p(v)$  and  $\rho(v)$ ,

$$\frac{\partial^2 \Omega}{\partial \rho \partial a} = \left( \frac{1 - F(v)}{f(v)} \right) \frac{\partial H(p(v) - v | \rho)}{\partial \rho} \Big|_{\rho=\rho(v)},$$

which is strictly negative when  $p(v) > v$  and strictly positive when  $p(v) < v$ . Hence, the set of values of  $\rho$  that maximizes  $\Omega$  is decreasing with  $a$  for  $v < p(v)$  and increasing with  $a$  for  $v > p(v)$ , strictly so when it contains a maximizer  $\rho > 0$  (cf. again Vives [21, Theorem 2.3]), which proves the asserted comparison.

*Incentive compatibility.* The final step in the proof is now to verify that the characterized solution to the relaxed program indeed satisfies incentive compatibility. That is, we show that for all  $(v, \hat{v})$  it holds that

$$u(v) = u(v, v) \geq u(v, \hat{v}). \tag{20}$$

Note first that from the single-crossing property that holds everywhere for  $p$  and piecewise over  $v \geq \tilde{v}$  and  $v \leq \tilde{v}$  for  $\rho$ , it follows from standard arguments, using monotonicity of  $p(v)$  and  $\rho(v)$ , that any type  $v \geq \tilde{v}$  ( $v \leq \tilde{v}$ ) has no incentive to deviate to any  $\hat{v} \geq \tilde{v}$  ( $\hat{v} \leq \tilde{v}$ ), where  $\hat{v} \neq v$ . We thus focus on the remaining case where such a deviation would be “across” the threshold  $\tilde{v}$ . For this we consider, without loss of generality for the argument, some type  $v > \hat{v}$  and, hence, that  $\hat{v} < \tilde{v} < v$ . Suppose further that  $\rho(\hat{v}) \leq \rho(v) < \rho(\tilde{v})$ . A fully analogous argument applies to

the case where  $\rho(v) < \rho(\hat{v}) < \rho(\tilde{v})$ . With these specifications, we can transform the requirement (20) to  $\Phi(v, \hat{v}) \geq 0$ , using

$$\begin{aligned} \Phi(v, \hat{v}) &:= u(v, v) - u(v, \hat{v}) \\ &= \int_{\hat{v}}^v \left[ \int_{p(\zeta)}^{p(\hat{v})} h(p - \zeta \mid \rho(\hat{v})) dp + \int_{\rho(\hat{v})}^{\rho(\zeta)} \left[ -\frac{\partial H(p(\zeta) - \zeta \mid \rho)}{\partial \rho} \right] d\rho \right] d\zeta. \end{aligned}$$

The first integral is clearly positive. Performing a change in the order of integration and after some transformations, the second integral becomes

$$\begin{aligned} &\int_{\rho(v)}^{\rho(\tilde{v})} \underbrace{[k'(\rho(\varphi_2(\rho))) - k'(\rho(\varphi_1(\rho)))]}_{=0} d\rho \\ &+ \int_{\rho(\hat{v})}^{\rho(v)} \underbrace{[k'(\rho(\varphi_2(\rho(v)))) - k'(\rho(\varphi_2(\rho)))]}_{\geq 0} d\rho \geq 0, \end{aligned} \tag{21}$$

where, using strict monotonicity of  $\rho(v)$  over the two intervals  $v \leq \tilde{v}$  and  $v \geq \tilde{v}$ ,  $\varphi_1(\rho)$  and  $\varphi_2(\rho)$  are the respective inverse functions of  $\rho$ , such that  $\varphi_1(\rho) \leq \varphi_2(\rho)$  holds in the presently analyzed case. (Observe that, as noted in the main text, we use for (21) that for all  $v$  and  $v'$  with  $\rho(v) = \rho(v')$  the marginal rate of substitution (between  $i$  and  $\rho$ ) is the same, when evaluated at this (optimal) value of  $\rho$ , which solves (12).)  $\square$

**Proof of Corollary 1.** Take first types  $v < v^{**}$ . For  $v^*$ , the characterization of  $i(v)$  in (19) simplifies to

$$i(v^*) = \int_{p(v^*)-v^*}^{\infty} [v^* + \tilde{s} - p(v^*)] dH(\tilde{s} \mid \rho(v^*)).$$

When  $v^*$  is interior, this together with  $\hat{\omega}(v^*) = 0$  obtains that  $i(v^*) = k(\rho(v^*))$ , as asserted. Note next that for all  $v < v^{**}$  we obtain from (19) that

$$i(v) - k(\rho(v)) = \hat{\omega}(v) - \int_{v^*}^v [1 - H(p(v') - v' \mid \rho(v'))] dv'.$$

As we can use from (6) that  $i(v) = i(v^*) + \int_{v^*}^v \frac{\partial i(v')}{\partial v'} dv'$ , the assertion on the seller's margin from selling information thus follows as, at points of differentiability,

$$\frac{d[i(v) - k(\rho(v))]}{dv} = -\frac{dp(v)}{dv} (1 - H(p(v) - v \mid \rho(v))) \geq 0. \quad \square$$

**Proof of Proposition 2.** To solve the relaxed program, we proceed by first characterizing the solution for the three possible cases. Then, we show uniqueness of the solution obtained.

*Characterization for the three cases.* As noted in the main text, the characterization for when the participation constraint is slack at  $\bar{v}$  is the same as in Proposition 1 (Case 1). Note next that

this case can only apply when  $v^* < m$ , while the participation constraint is then indeed satisfied everywhere (and strictly so at  $\bar{v}$ ) if

$$\int_{v^*}^{\bar{v}} [1 - H(p(v) - v \mid \rho(v))] dv > \bar{v} - m. \tag{22}$$

In Case 2, where we assume that now  $u(v^*) > w(v^*)$  and  $u(\bar{v}) = w(\bar{v})$  (implying  $v^* = \underline{v}$ ), we can use, as in (6), that

$$u(v) = w(\bar{v}) - \int_v^{\bar{v}} [1 - H(p(v') - v' \mid \rho(v'))] dv',$$

such that after substitution into  $\Pi$  we have  $\Pi = \int_{v^*}^{\bar{v}} \hat{w}(v) f(v) dv - w(\bar{v})$ , where now  $\hat{w}(v)$  is given by (15) with  $\kappa = F(v^*)$ . Pointwise maximization uniquely obtains  $p(v)$  as in (17) for  $\kappa = F(v^*)$ , while after substitution of  $p(v)$ , when  $\rho(v) > 0$  holds, then this must still solve (12). Note now that the definition of  $\bar{v}$  clearly extends to any  $p(v)$  as given by (17) (i.e., to any feasible  $\kappa$ ), while it is strictly decreasing in  $\kappa$ . The participation constraint then holds everywhere (and strictly so at  $v^*$ ) if

$$\int_{v^*}^{\bar{v}} [1 - H(p(v) - v \mid \rho(v))] dv < \bar{v} - m. \tag{23}$$

In the third and final case, where the constraint binds at both  $v^*$  and  $\bar{v}$ , we can use the approach in Esö and Szentes [7] to avoid the use of optimal control techniques (and, thereby, making restrictions on  $\rho(v)$ ). By this procedure, one supposes that the seller could observe whether  $v < v_0$  or  $v > v_0$  for some  $v^* \leq v_0 \leq \bar{v}$ , thus solving the two problems where for  $v < v_0$  (and the respective conditional distribution for  $F(v)$ ) only the constraint  $u(v^*) \geq w(v^*)$  is considered and where for  $v > v_0$  (and the respective conditional distribution for  $F(v)$ ) only the constraint  $u(\bar{v}) \geq w(\bar{v})$  is considered. When one can ultimately choose a value  $v^* \leq v_0 \leq \bar{v}$  such that  $u(v)$  is continuous at  $v_0$ , the joint solution for the two problems also solves the original (relaxed) problem. Proceeding in this way, though for brevity's sake referring to Esö and Szentes [7] for the details, for Case 3 this obtains a value  $\kappa$  satisfying  $F(v^*) \leq \kappa \leq 1$ , with  $p(v)$  solving (17) and, given  $p(v)$ ,  $\rho(v) > 0$  solving (12). When such a construction is feasible, then for given  $\kappa$  it holds that

$$\int_{v^*}^{\bar{v}} [1 - H(p(v) - v \mid \rho(v))] dv = \bar{v} - m. \tag{24}$$

Note now that here and in what follows we omit the characterization of  $i(v)$ , as well as that of  $q$  (together with  $v^{**}$ ). In analogy to the proof of Proposition 1, this follows immediately, once  $p(v)$  and  $\rho(v)$  are characterized. Furthermore, the results on global incentive compatibility from Proposition 1 clearly remain valid, irrespective of which of the Cases 1–3 apply.<sup>35</sup>

<sup>35</sup> As already noted above, holding all else constant, Case 1 applies for all sufficiently high values  $m$ . To see that also Cases 2 and 3 are not purely hypothetical, consider the extreme case where  $\underline{v} = c = m$  (or, more generally,  $\underline{v} + \varepsilon = c =$

Unique determination of  $\kappa$  and  $v^*$  satisfying  $\underline{v} \leq v^* < \tilde{v} < v^{**} \leq \bar{v}$ . As a next step, we show that the characterization of  $F(v^*) \leq \kappa \leq 1$  is unique. To see this, observe first that as  $\kappa$  increases, where  $\kappa = 1$  and  $\kappa = F(v^*)$  comprise the first and the second case, then  $p(v)$  increases for all  $v$ . (Recall again that we extend the definition of  $p(v)$  to all  $v \in [v^*, \bar{v}]$ , i.e., even when  $\rho(v) = 0$ .) Next, from inspection of (12) we have that, when considering a marginal increase in  $\kappa$  and comparing the respective optimal values  $\rho(v)$ , then the optimal value  $\rho(v)$  after the increase in  $\kappa$  is higher when  $v > \tilde{v}$  and lower when  $v < \tilde{v}$  (both strictly when  $\rho(v) > 0$  holds either before or after the change in  $\kappa$ ).<sup>36</sup> Taking together the changes in  $p(v)$  and  $\rho(v)$  that result from a marginal increase in  $\kappa$ , we thus have that over all  $v \in [v^*, \bar{v}]$  the slope  $du/dv = 1 - H(p(v) - v | \rho(v))$  (when  $u(v)$  is differentiable) must decrease (and strictly so when  $\rho(v) > 0$  holds either before or after the change). Consequently, there can only exist one value  $\kappa$  where either one of the three conditions (22), (23), and (24) holds. (That such a value  $\kappa$  exists follows immediately from the obtained fact that  $\int_{v^*}^{\tilde{v}} [1 - H(p(v) - v | \rho(v))] dv$  is strictly decreasing in  $\kappa$ .)

Note next that at  $v = \tilde{v}$  we always have that the marginal value of information is constant and given by  $\frac{\partial}{\partial \rho} \int_0^\infty \tilde{s} dH(\tilde{s} | \rho)$ , which from (13) and by continuous differentiability of  $k(\rho)$  implies that  $\rho(\tilde{v}) > 0$ . As argued already in the proof of Proposition 1, this implies that  $\underline{v} \leq v^* < \tilde{v} < \bar{v}$  must hold, irrespective of which case applies and thus irrespective of  $\kappa \in [0, 1]$ . (Note also that  $\underline{v} < \tilde{v} < \bar{v}$  holds from  $\underline{v} < c < \bar{v}$ .)

To finally see that  $v^*$  and  $\kappa$  are jointly uniquely determined, note first that when Cases 1 or 2 apply this is immediate. For Case 3, note next that when  $v^*$  is interior, then  $v^*$  and  $\kappa$  must jointly solve (24) and

$$\int_{p(v^*)-v^*}^\infty \left[ v^* + \tilde{s} - c - \frac{\kappa - F(v^*)}{f(v^*)} \right] dH(\tilde{s} | \rho(v^*)) - k(\rho(v^*)) = 0. \tag{25}$$

Given that (24) defines a strictly decreasing and continuous relationship between  $v^*$  and  $\kappa$  and given that (25) defines an increasing and continuous relationship, this pins down a unique solution also for Case 3 (provided that such a solution exists, given that Case 3 applies). Taken together, we thus have that only one of the three cases can apply and that this then always pins down a unique solution  $v^*$  and  $\kappa$ . □

**Proof of Corollary 2.** A change of  $m$  has no impact on the characterization in Cases 1 and 2. In Case 3, note that a higher  $m$  decreases the right-hand side in (24). Note also that the left-hand side of (24) is strictly decreasing in  $v^*$  and  $\kappa$ . Recall next that, from (25),  $v^*$  has to increase with  $\kappa$  when  $v^*$  is interior. Taken together, as  $m$  increases, we have in Case 3 that  $\kappa$  must strictly decrease, while  $v^*$  strictly decreases when interior and may, otherwise, stay constant at  $v^* = \underline{v}$ . As  $m$  increases, we also have from (22)–(24) that any transition must be from Case 1 to Case 3, and then to Case 2. From (17), we have immediately that as  $\kappa$  decreases, following a reduction in  $m$ , this reduces all prices  $p(v)$ . As the price without information,  $q$ , is equal to  $m$ , when  $u(\bar{v}) = w(\bar{v})$ , it follows that the seller’s margin for the good decreases for all  $v$ . (In fact, it even becomes negative for  $v_0 < v < v^{**}$ .) This completes the proof of the asserted comparative result in (i).

$m - \varepsilon$  for  $\varepsilon > 0$  sufficiently small). We can show that then Case 2 must apply, with  $v^* = \underline{v}$ . (Formally, this holds as condition (23) is then surely satisfied at  $v^* = \underline{v}$ .) Holding everything else constant, as we now increase  $m$ , we thus indeed “walk through” all three cases (cf. also the comparative statics results below).

<sup>36</sup> Considering only a marginal change allows to abstract from the joint change in  $\tilde{v}$ , given that this is continuous in  $\kappa$ .

For the comparative analysis in the quality of information, we first obtain for  $\rho > 0$  that

$$\frac{\partial^2 \hat{\omega}}{\partial \rho \partial \kappa} = \frac{\partial p(v)}{\partial \kappa} \frac{\partial H(p(v) - v | \rho)}{\partial \rho} \Big|_{\rho=\rho(v)}, \tag{26}$$

where it holds from (10) that  $\partial p(v)/\partial \kappa > 0$ . Hence, when  $v < \tilde{v}$ , using (1) and also that  $\tilde{v}$  is continuous in  $\kappa$ , we have from standard monotone comparative statistics results (cf. again Vives [21, Theorem 2.3]) that any selection  $\rho(v)$  must be decreasing in  $\kappa$  (and strictly so when previously  $\rho(v) > 0$ ), while when  $v > \tilde{v}$  then it must be increasing (and strictly so when this results in  $\rho(v) > 0$ ). From this observation together with strict monotonicity of  $\tilde{v}$  in  $\kappa$ , we immediately get the following. Denote explicitly the values  $\tilde{v}(\kappa = 0)$  and  $\tilde{v}(\kappa = 1)$ , which satisfy  $\underline{v} < \tilde{v}(\kappa = 0) < \tilde{v}(\kappa = 1) < \bar{v}$ . Now, as  $\kappa$  decreases, as the substitute good becomes more attractive (lower  $m$ ), it holds that (provided the respective types participate), when  $v < \tilde{v}(\kappa = 0)$ ,  $\rho(v)$  is always increasing, while for  $v > \tilde{v}(\kappa = 1)$  it is always decreasing. For all  $\tilde{v}(\kappa = 0) < v < \tilde{v}(\kappa = 1)$ , there is a value  $0 < \tilde{\kappa} < 1$  such that as  $\kappa$  decreases from  $\kappa = 1$  to  $\kappa = 0$ , then  $\rho(v)$  first increases, as long as still  $\kappa > \tilde{\kappa}$ , and then decreases, when  $\kappa < \tilde{\kappa}$ , completing the proof of assertion (ii) in the corollary.  $\square$

**Proof of Corollary 3.** We again only consider the case where  $m \geq c_G$ . Note first that with  $F(v_0) = \kappa$  and thus  $\tilde{v} = c + \frac{F(v_0) - F(\tilde{v})}{f(\tilde{v})}$ , in Case 1, where  $\kappa = 1$ , it holds that  $v_0 = \bar{v} > \tilde{v}$ , while in Case 2, where  $\kappa = 0$ , it must hold that  $v_0 = v^* = \underline{v} < \tilde{v}$ . Further, both  $v_0$  as well as  $\tilde{v}$  are increasing in  $\kappa$  with

$$\begin{aligned} \frac{dv_0}{d\kappa} &= \frac{1}{f(v_0)}, \\ \frac{d\tilde{v}}{d\kappa} &= \frac{1}{f(\tilde{v})} \frac{1}{1 - p'(\tilde{v})} < \frac{1}{f(\tilde{v})}. \end{aligned}$$

From the last observation there exists some  $0 < \kappa' < 1$  such that  $v_0 > \tilde{v}$  when  $\kappa > \kappa'$  and  $v_0 < \tilde{v}$  when  $\kappa < \kappa'$ , while at  $\kappa'$  it must hold that  $v_0 = \tilde{v} = c$  (or  $\kappa' = F(c)$ ). For  $\kappa = \kappa'$  the type  $v = v_0 = \tilde{v}$  receives both the efficient price  $p(v_0) = c$  and the efficient level of information. In general, type  $\tilde{v}$  always receives the efficient level of information and further  $\rho(\tilde{v}) > 0$ , while type  $v_0$  always receives the efficient price  $p(v_0) = c$  (only pinned down uniquely in case  $\rho > 0$ ) and from  $\hat{\omega}(v_0) = \omega(v_0)$  also the efficient level of information.

Next, again consider  $\Omega(\rho(v), a) = \omega(v) - a[\hat{\omega}(v) - \omega(v)]$  and observe that now

$$\frac{\partial^2 \Omega}{\partial \rho \partial a} = \frac{\kappa - F(v)}{f(v)} \frac{\partial H(p(v) - v | \rho)}{\partial \rho} \Big|_{\rho=\rho(v)},$$

which is positive if either  $v < v_0$  and  $v > \tilde{v}$  or if  $v > v_0$  and  $v < \tilde{v}$ , zero if either  $v = v_0$  or  $v = \tilde{v}$ , and, finally, negative if either  $v < v_0$  and  $v < \tilde{v}$  or  $v > v_0$  and  $v > \tilde{v}$ . The under- or overprovision results for the different cases then follow immediately.  $\square$

### Appendix B. General mechanism

As noted in the main text, we now show that the analyzed simple mechanism implements the outcome under the optimal general mechanism. To prove this, we follow Esö and Szentes [6,7] and first suppose that also the posterior valuation  $\tilde{s}$  was verifiable. As the outcome of an optimal mechanism in this case is also implemented by the mechanism characterized in Proposition 1,

offering the respective simple mechanism is indeed optimal for the seller. As the proof thus closely follows Esö and Szentes [6,7], we will be brief.

Without loss of generality, using risk neutrality of the seller and the buyer, we can specify that the considered, general mechanism prescribes the following: (i) An expected transfer  $r(v)$ ; (ii) a possible random choice of information quality, as expressed by some distribution  $G(\rho | v)$  with support  $\varrho(v)$ ; (iii) and a probability of trade  $\gamma(\tilde{s}, \rho, v)$ , depending on the subsequently realized value  $\tilde{s}$ .

We use again that from incentive compatibility types  $v \in [v^*, \bar{v}]$  participate. With this mechanism, the utility of buyer  $v$  when reporting  $\hat{v}$  becomes

$$u(v, \hat{v}) = \int_{\varrho(\hat{v})} \left[ \int_{-\infty}^{\infty} \gamma(\tilde{s}, \rho, \hat{v}) [v + \tilde{s}] dH(\tilde{s} | \rho) \right] dG(\rho | \hat{v}) - r(\hat{v}),$$

while under truthtelling with  $u(v) = u(v, v)$ , the seller’s expected profits from type  $v$  equals

$$\pi(v) := \int_{\varrho(v)} \left[ \int_{-\infty}^{\infty} \gamma(\tilde{s}, \rho, v) [v + \tilde{s} - c_G] dH(\tilde{s} | \rho) - k(\rho) \right] dG(\rho | v) - u(v).$$

Using that  $u(v^*) = 0$  holds by optimality, we obtain from incentive compatibility

$$u(v) = \int_{v^*}^v \left[ \int_{\varrho(v')} \left[ \int_{-\infty}^{\infty} \gamma(\tilde{s}, \rho, v') dH(\tilde{s} | \rho) \right] dG(\rho | v') \right] dv',$$

such that the seller’s profits become, after substitution for  $u(v)$  and partial integration,

$$\begin{aligned} \Pi &= \int_{v^*}^{\bar{v}} \left[ \int_{\varrho(v)} \left[ \int_{-\infty}^{\infty} \gamma(\tilde{s}, \rho, v) \right. \right. \\ &\quad \times \left. \left. \left[ v + \tilde{s} - c_G - \frac{1 - F(v)}{f(v)} \right] dH(\tilde{s} | \rho) - k(\rho) \right] dG(\rho | v) \right] dF(v), \end{aligned} \tag{27}$$

which we can maximize pointwise for all  $v \in [v^*, \bar{v}]$ . We show that for all  $v$  the simple mechanism characterized in Proposition 1 implements a solution to the general mechanism.

Denote  $p(v) = c_G + \frac{1 - F(v)}{f(v)}$ , such that, for given  $v$ , any  $\rho \in \varrho(v)$  maximizes, together with a corresponding optimal choice of  $\gamma(\tilde{s}, \rho, v)$ , the objective

$$\int_{-\infty}^{\infty} \gamma(\tilde{s}, \rho, v) [v + \tilde{s} - p(v)] dH(\tilde{s} | \rho) - k(\rho). \tag{28}$$

Note that, for given  $\rho$ , optimality thus requires to set  $\gamma(\tilde{s}, \rho, v) = 1$  when  $v + \tilde{s} - p(v) > 0$  and to set  $\gamma(\tilde{s}, \rho, v) = 0$  when  $v + \tilde{s} - p(v) < 0$ . This corresponds to the buyer’s optimal decision, when  $p(v)$  is the price that he must pay at the purchasing stage. Moreover, any choice of  $\rho(v)$  as in Proposition 1 also satisfies  $\rho \in \varrho(v)$ , while specifying that  $\rho(v) = \varrho(v)$  is also optimal in the general mechanism (and vice versa). (When  $\varrho(v)$  is not a singleton, however, then also randomization would be optimal.)

Note next that the payments  $r(v)$  are obtained from  $u(v^*) = 0$  next to the requirement that, at points of differentiability,  $du/dv = \partial u/\partial v$ , which given the obtained results can again be written as  $\partial u/\partial v = 1 - H(p(v) - v | \rho(v))$  also for the general mechanism. The obtained function  $r(v)$  can then be implemented in the simple mechanism by choosing, when  $\rho(v) > 0$ ,  $i(v) = r(v) - p(v)[1 - H(p(v) - v | \rho(v))]$ , while otherwise setting  $q = r(v)$ .

## References

- [1] A. Araujo, H. Moreira, Adverse selection problems without the Spence–Mirrlees condition, *J. Econ. Theory* 145 (2010) 1113–1141.
- [2] D. Baron, D. Besanko, Regulation and information in a continuing relationship, *Info. Econ. Pol.* 1 (1984) 267–302.
- [3] H. Bar-Isaac, G. Caruana, V. Cufinat, Information gathering and marketing, *J. Econ. Manage. Strategy* 19 (2008) 375–401.
- [4] G. Biglaiser, C. Mezzetti, Principals competing for an agent in the presence of adverse selection and moral hazard, *J. Econ. Theory* 61 (1993) 309–330.
- [5] P. Courty, H. Li, Sequential screening, *Rev. Econ. Stud.* 67 (2000) 697–717.
- [6] P. Esö, B. Szentes, Optimal information disclosure in auctions and the handicap auction, *Rev. Econ. Stud.* 74 (2007) 705–731.
- [7] P. Esö, B. Szentes, The price of advice, *RAND J. Econ.* 38 (2007) 863–880.
- [8] J. Ganuza, J. Penalva, Signal orderings based on dispersion and the supply of private information in auctions, *Econometrica* 78 (2010) 1007–1030.
- [9] F. Hoffmann, R. Inderst, Price discrimination and the provision of information, Working Paper, Goethe University Frankfurt, 2009.
- [10] F. Hoffmann, R. Inderst, Supplementary material to pre-sale information, available at <http://www.florian-hoffmann.com/research>, 2011.
- [11] B. Jullien, Participation constraints in adverse selection models, *J. Econ. Theory* 93 (2000) 1–47.
- [12] J. Johnson, D. Myatt, On the simple economics of advertising, marketing, and product design, *Amer. Econ. Rev.* 96 (2006) 756–784.
- [13] D. Kräbmer, R. Strausz, Optimal procurement contracts with pre-project planning, *Rev. Econ. Stud.* 78 (2011) 1015–1041.
- [14] T. Lewis, D. Sappington, Supplying information to facilitate price discrimination, *Int. Econ. Rev.* 35 (1994) 309–327.
- [15] T. Lewis, D. Sappington, Information management in incentive problems, *J. Polit. Economy* 105 (1997) 796–821.
- [16] G. Maggi, A. Rodriguez-Clare, On countervailing incentives, *J. Econ. Theory* 66 (1995) 238–263.
- [17] P. Milgrom, I. Segal, Envelope theorems for arbitrary choice sets, *Econometrica* 70 (2002) 583–601.
- [18] Oxera, Retail distribution review proposals: impact on market structure and competition, Report prepared for the Financial Services Authority, 2009.
- [19] A. Saak, The value of private information in monopoly, *J. Ind. Econ.* 56 (2008) 169–189.
- [20] C. Terwiesch, C. Loch, Collaborative prototyping and the pricing of custom-design products, *Manage. Sci.* 50 (2004) 145–158.
- [21] X. Vives, *Oligopoly Pricing: Old Ideas and New Tools*, MIT Press, Cambridge, MA, 2000.