A Mechanism Design Approach to the Optimal Disclosure of Private Client Data

Job Market Paper

Find the latest version at
www.tokiskostas.com/JMpaper.pdf

Konstantinos Tokis*

November 26, 2017

Abstract

This paper studies the incentives of a seller to voluntarily disclose or sell information about a buyer to a third party. While there are obvious benefits to sharing information with other sellers, there is also an incentive cost which is due to her learning about the buyer through her own trade with him. To study this trade-off we analyse a model in which a buyer interacts sequentially with two sellers, each of whom makes a take-it-or-leave-it offer. The buyer learns his valuation for the good of each seller sequentially but these might be correlated. In addition, we model information disclosure using Bayesian persuasion, that is we allow the first seller to commit to a disclosure rule which depends on the information she acquires in the first trade. In this setting we fully characterise the first seller's costs and benefits of information sharing. In particular, we show that voluntary information disclosure, or selling of information, is optimal when the correlation between the buyer's valuations for the two goods is not too positively correlated. Also, when information exchange is optimal the buyer benefits from it if his valuations are positively correlated, otherwise he is worse off.

Keywords: Mechanism Design, Information Design, Disclosure, Bayesian Persuasion, Sequential common agency games

JEL codes: D82, D86, C73, L1

*Ph.D. Candidate in Economics, Department of Economics, The London School of Economics and Political Science, E-mail: k.e.tokis@lse.ac.uk, Website: http://www.tokiskostas.com/. I am indebted to my advisor Ronny Razin for his continuous support and guidance. This project would not have being possible without the generous financial support of the Paul Woolley Centre for the Study of Capital Market Dysfunctionality. I would also like to thank Dimitri Vayanos, Gilat Levy, Balázs Szentes, Erik Eyster, Philippe Aghion, Francesco Nava, Matthew Levy, Andrew Ellis, Clement Minaudier, Alex Moore, Dimitris Papadimitriou, George Vichos, Panos Mavrokonstantis, Alkiviadis Georgiadis-Harris, and all the participants of the LSE Micro Theory group for helpful comments and discussions.

1 Introduction

Information exchange between firms is rapidly increasing in both volume and importance. The introduction of new technologies allows firms to cheaply store, analyse, and share data on their clients. In particular, a firm just by interacting with its customers acquires private information on their preferences. This information is valuable to other firms. To give an example, an architect learns her client's willingness to pay to renovate his house. This is valuable information for an interior designer, since the client's preferences for the two services should be highly correlated. By sharing information with the designer the architect might be able to increase her profits. More generally, information exchange is significant in business to business environments. Notably, data brokerage is already a big industry and has recently attracted a lot of attention from both firms, that look to expand their customer base, and regulators[1].

This paper poses the following question: "Should a firm share private information about its customers?" One direct way to benefit from disclosing data on its clientèle, such as their purchase history, is to sell it. But even if selling information is not an option a firm can still benefit from disclosure, but in an indirect way. This is because selectively sharing data on its clients could persuade other firms to offer them discounts, the prospect of which would allow the firm to increase its own prices. To see this, suppose that the client of the aforementioned architect opted for a partial renovation only, which was the cheapest option. Then he probably doesn't value that much the interior designer's services. Therefore, by disclosing this information the architect can persuade the designer to offer a discount to her client. In turn this allows the architect to sell to her client both the partial renovation and the prospect of a discount as a bundle of products.

However, information disclosure is associated with an incentive cost. This is because a firm infers its customers' willingness to pay indirectly from their choice of product within its catalogue. But if those are aware that this choice may affect the probability of their getting a discount from a related seller, then this will skew their choices towards cheaper products. In the example above, a client who anticipates a possible discount may skew his choice towards the partial renovation. Therefore, information provision is interwoven with an incentive cost for acquiring it.

To answer the question of optimal disclosure we use a mechanism design approach coupled with Bayesian persuasion. We consider a two period model in which two sellers sequentially interact with a single buyer. The first and second seller make take-it-or-leave-it offers to

[1] Articles on the data brokerage industry have recently being published by The Economic Times, The Atlantic, and The NY Times among others. The predominant legislative change is the EU General Data Protection Regulation.

the buyer in the first and second period, respectively. Each seller can offer one unit of an indivisible good. The buyer's valuation for each of the two goods is either high or low, and evolves stochastically between the two trades. That is when trading with the first seller the buyer does not know his valuation of the good of the second seller. Despite that, the buyer's valuations of the two goods are correlated, so his first period preferences are valuable information to the second seller.

We model information disclosure by using a Bayesian persuasion framework. To be more precise, we assume that the first seller can commit ex ante to a distribution of signals that depends on the buyer's report in her mechanism. The realisation of this signal is observed by the second seller, who uses this information when determining her price.

At this point it is helpful to first consider the case of perfect positive correlation, which implies that the buyer always assigns the same value to the goods of the first and second seller. Hence while trading at period 1, the buyer faces no uncertainty over his valuation of the second good. We show that in this case non-disclosure is always optimal, a result that has been previously stated by Calzolari and Pavan (2006). The reason for the optimality of non-disclosure is twofold. The first is that a period 1 low type buyer assigns no value to the potential second period discount, as his period 2 type will also be low. Hence he obtains zero rents from this trade. The second reason is that the incentive cost of convincing the period 1 high type to be truthful is the highest possible. This is because he knows with certainty that his second period valuation will be high, thus this is when he values the potential discount the most. Consider now the implications of the above intuition for the case of imperfect correlation. It hints that if we increase the probability of a period 1 low type to become high, and decrease the probability of a period 1 high type to remain high, then information disclosure might become optimal.

The main result of our paper is that indeed when the buyer is uncertain about his future valuations information disclosure is optimal for a substantial set of environments. In particular, this uncertainty is a result of imperfect correlation between the valuations of the two sellers. In Propositions 2 and 3 we characterise all the correlation structures of the buyer's valuations for which information disclosure is optimal. In addition, we derive the corresponding optimal disclosure policy.

We start our analysis with the first best problem of the first seller. In this case the buyer's period 1 valuation is observable by the first seller, but not by the second. This analysis allows us to characterise the benefit of information disclosure abstracting away from any incentive costs. Proposition 1 characterises the set of environments for which this benefit is strictly positive. The optimal disclosure policy takes a simple form. When the buyer's types are positively correlated this signal randomises between revealing the high type and pooling it together with the low type. The pooling outcome is the one that creates a discount as discussed in the example above, which provides the benefit of information disclosure. Diametrically, under negative correlation it is the low type that is some times revealed (shown in Proposition 3, which focuses on negative correlation).

We next analyse the second best in which the first seller incurs incentive costs for eliciting the buyer's valuation for her good. Proposition 2 characterises the set of positive correlation structures for which information disclosure is optimal and shows that this is a non-empty convex subset of the corresponding set characterised in the first best (Proposition 1). As we show below this set sometimes includes correlation structures that are arbitrarily close to perfect positive correlation. Proposition 3 provides a similar analysis for the case of negative correlation showing that disclosure is optimal whenever it is optimal under the first best. In addition, we show that the optimal disclosure policy, whenever disclosure is optimal, is always the same as in the first best for both positive and negative correlation structures.

We also demonstrate that the buyer might both benefit or lose from information disclosure. While the low types are always indifferent, it is the high types that are influenced by disclosure. In the case of positive correlation, disclosure opens up the possibility of discounts, which are beneficial for the high type buyer. On the other hand, in the negative correlation case, when information disclosure is optimal the optimal mechanism involves lowering the rents of the high type to increase profits.

On a more technical note our analysis proceeds on the following way. First, we derive the buyer's payoff from his second trade. This is positive only if (i) the buyer is a high period 2 type and (ii) the posterior of the second seller is low enough for her to choose to sell to both types. Hence while trading with the first seller the buyer's expected payoff from the second period is a function of (i) the distribution of signals which corresponds to the reported type and (ii) the probability of being a high period 2 type. Thus, when the agent's types are positively correlated information provision will result in additional rents for the high period 1 type, since the low type's distribution will necessarily entail a more frequent discount. We use insights from Gentzkow and Kamenica (2011) to reformulate the seller's information provision problem as a choice of distributions over posteriors. This allows us to use a graphical solution that relies on the concave closure of our objective function. Finally, we find two type depended signal distributions that implement the optimal unconditional distribution of posteriors.

We extend the analysis in several ways. First, we consider the case when the first seller can sell her information to the second seller. We show that the solution of this model is identical to that of the baseline one. In an alternative extension we allow the two sellers to produce a continuous quantity under an isoelastic cost function, which results in a solution comparable to that of the baseline model. In the online appendix we further extend this model to multi-period contracts, we show how those can incorporate moral hazard, and demonstrate the interplay between information provision and endogenous termination times in this setting. Finally, we provide some sufficient results for non-disclosure to be optimal under continuous types.

My paper is closely related to Calzolari and Pavan (2006), the implications of which the authors have further explored in Calzolari and Pavan (2008, 2009). They also consider a setup where an agent sequentially contracts with two principals, the first of whom can commit on a disclosure rule. However, they restrict their attention to the cases of perfect positive and negative correlation. The implication of this is that at the point of trading with the first seller the buyer knows both his valuations. Under perfect positive correlation, they obtain that privacy is always optimal. However, under perfect negative correlation, and when some additional conditions are satisfied, they establish that the first seller's optimal signal could be informative. In addition, the authors show that an alternative way to make disclosure optimal is to assume externalities. In this paper we examine a conceptually similar model while using a Bayesian persuasion framework, which is a natural setting to consider the case of imperfect correlation. We aim to demonstrate that information exchange can arise under much more natural and economically interesting conditions. Our analysis relies neither on externalities, nor on the extreme assumption of perfect negative correlation. Notably, we show that even close to perfect positive correlation could be enough for some disclosure to be optimal, which implies that the aforementioned result on the optimality of privacy is not robust to stochastic preferences.

Another very relevant paper is that of Dworczak (2016*a*), where a seller auctions an object and the winner's payoff depends on both his type and a generic aftermarket. Crucially, the aftermarket related payoff depends on the public posterior on the winner's type. This introduces an information design aspect to the auctioneer's problem. The author identifies a class of mechanisms, called cutoff mechanism, that are implementable regardless of the aftermarket's form. He subsequently identifies the optimal mechanism within this class, which for the single agent model has no information disclosure. On an accompanying paper Dworczak (2016*b*), provides sufficient condition for a cutoff mechanism to be optimal. Among other results he shows that if the seller only acts as an information intermediary, that is she buys information from the bidders and sells it to the aftermarket, then information provision is never implementable.

Our paper presents an interplay of mechanism and information design. In some sense it is related to the literature of dynamic mechanism design, since we allow for the buyer's preferences to evolve stochastically over time (for example see Pavan et al. (2014); Garrett and Pavan (2012); Eső and Szentes (2017) and the references therein). In particular, the dynamic extension of our model, which can be found on the online appendix, follows closely Battaglini (2005). He considers a firm contracting with a consumer, whose private valuation of the firm's product is either high or low and evolves stochastically according to an exogenous Markov matrix.

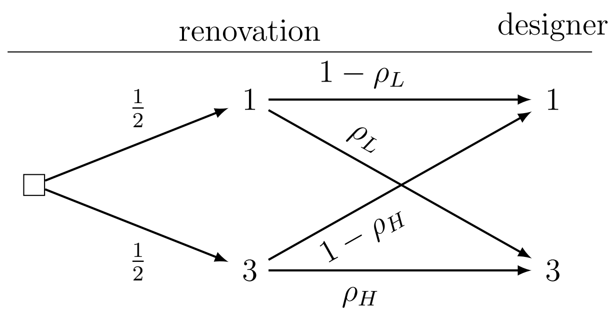
Nevertheless, in contrast to this literature in our model the first seller can only indirectly affect the choices of the second period, that is of seller two. To be more precise, her only tool is the disclosure policy she commits on. Hence the first seller engages in Bayesian persuasion with the second on the buyer's reported type. In that sense our analysis is related to the literature of information design (for example see Inostroza and Pavan (2017) and the references therein). In particular, the solution method that we use is related to the work of Gentzkow and Kamenica (2011), who provide a framework for solving a vast class of information design problems. In a more recent paper, Ely (2017) considers a dynamic information design problem where the designer receives signals on the underline state overtime. Roesler and Szentes (2017) revisit the canonical bilateral trade model. They assume that the buyer is uncertain about her valuation of the product, but receives a signal on it. They derive the buyer-optimal distribution of signals and show that this generates efficient trade.

The rest of the paper proceeds as follows. Section 2 provides an example that further clarifies the potential benefits and costs of information provision, and how those depend on the stochasticity of the buyer's preferences. Section 3 formally defines our baseline model. Section 4 provides the analysis and the corresponding results. Section 5 considers two natural extensions of our model. The first allows for the seller to sell information, while the second for the buyer to know both his types at the first period. Finally, section 6 discusses the implications of the model and concludes.

2 Example

To fix ideas consider the following illustrative example. An architect makes a take-it-or-leave-it offer to a buyer for the renovation of his house. Suppose the architect has zero cost of production and that a high type buyer values the architect's services at $\theta_H = 3$, while a low type at $\theta_L = 1$. Assume throughout that the buyer's type (call it renovation type) is his private information, and that its two realisations are equiprobable. Standard argumentation shows that the architect's optimal pricing strategy is to only sell to the high type by setting price $\overline{p} = 3$.

Next we expand the space of interactions by assuming that the buyer can not only hire an architect, but also an interior designer. The latter will also make a take-it-or-leave-it offer to the buyer, but only after the architect has made hers, and has zero production cost. Despite that, the buyer can opt to hire the interior designer without hiring the architect and visa versa. Again, we set the buyer's valuation of the designer's services (call it design type) to be either $\theta_H = 3$ or $\theta_L = 1$. The buyer's valuations for the two services will not necessarily be the same, and even he will only learn the latter when the interior designer presents her product to him. Thus the buyer's uncertainty over his preferences is resolved sequentially. In particular, assume that a high renovation type remains a high design type with probability ρ_H, whereas a low renovation type turns into a high design type with probability ρ_L. The architect is able to connect the buyer with the interior designer. For simplicity, we assume that in the absence of such a connection the designer's posterior on the buyer's design type is $(\rho_H + \rho_L)/2$, which equals the public prior.



Through her interaction with the buyer, the architect acquires valuable private information on his preferences. This is because the buyer's willingness to pay $\overline{p} = 3$ for the renovation reveals him as a high type. This information is valuable to the designer because the buyer's renovation and designer types are correlated. Hence the designer could use the buyer's purchase history to decide between offering the high and low price, that is $\overline{p} = 3$ and $\underline{p} = 1$, respectively.

This begs the question "Is it possible for the architect to increase her expected revenue by using her private information on the buyer's preferences?". It turns out that the answer depends a lot on the degree of uncertainty that the buyer has on his own preferences, that is ρ_L and ρ_H. To demonstrate this suppose

$$\frac{\rho_H + \rho_L}{2} > \frac{1}{3}$$

Since $\theta_L/\theta_H = 1/3$ we infer that the designer's optimal pricing policy, in the absence of any communication with the architect, is to set $\overline{p} = 3$ and only sell to the high design type. Thus there is some room for the architect to attempt to persuade the designer, in the Bayesian sense, to offer a discount to the buyer. This could ultimately be beneficial for the architect, as she could charge the buyer for this discount[2].

To demonstrate this fix $(\rho_H, \rho_L) = (1/2, 1/4)$ and suppose that the architect is able to commit in advance to both the buyer and the interior designer that she will reveal a high renovation type with probability $g_H = 1/2$. Henceforth, whenever she sends a buyer to the designer without revealing him as a high renovation type the latter's posterior on the former's design type to be high is

$$\frac{1 - g_H}{1 + (1 - g_H)} \cdot (\rho_H - \rho_L) + \rho_L \; = \; \frac{1}{3}$$

Thus, not revealing the buyer as a high renovation type achieves a discount, since the designer's optimal price becomes $\underline{p} = 1$. The benefit of this signal for the architect is that she is able to charge the low house type for this discount with price

$$p_L \; = \; \rho_L \cdot (\overline{p} - \underline{p}) \; = \; \rho_L \cdot (\theta_H - \theta_L)$$

whereas when she was not engaging in bayesian persuasion her interaction with the low renovation type had zero value, since such a type would not buy from her at her previously optimal price $\overline{p} = 3$. It is worth mentioning that this benefit exists only under imperfect correlation as otherwise $\rho_L = 0$.

However, this informative signal generates an incentive cost[3], which will decrease the price that the architect can charge to the high renovation type p_H below θ_H. This is because the high type is tempted to behave like a low type since in this case he gets the design discount with probability 1, instead of $1 - g_H = 1/2$. Hence the maximum difference between p_H and p_L has to be $\theta_H - \rho_H(\theta_H - \theta_L)/2$, otherwise even the high type would opt to get the guaranteed design discount instead of the uncertain one together with the renovation. As a result, the usage of the signal will decrease the price charged to the high renovation type by

$$\frac{\rho_H}{2}(\theta_H - \theta_L) - p_L \; = \; \left(\frac{\rho_H}{2} - \rho_L\right)(\theta_H - \theta_L)\,,$$

Interestingly, the benefit of information provision is associated to the architect's trade with the low renovation type, whereas the cost to her trade with the high type. Therefore to find the net impact that the usage of the informative signal has on the architect's revenue

[2]Note that the architect could not be benefitted by just establishing a connection and not providing any information on the buyer's preferences. In addition, this would be true even if the buyer could only hire the interior designer through the architect. This is because the net benefit of this connection for the buyer would be zero, since a low design type will not buy at $\overline{p} = 3$ and even a high type will be indifferent between buying or not.

[3]The discussion here ignores the incentive compatibility constrain of the low type, but it is easy to check that it holds in all the cases that are considered.

multiply both the benefit and the cost by 1/2, which is the probability of facing each type, and subtract the latter from the former

$$\frac{1}{2}\,\rho_L(\theta_H-\theta_L)-\frac{1}{2}\left(\frac{\rho_H}{2}-\rho_L\right)(\theta_H-\theta_L) \;=\; \left(\rho_L-\frac{\rho_H}{4}\right)(\theta_H-\theta_L)$$

But under the assumed transitioning probabilities $\rho_L - \rho_H/4 = 1/16$, which is positive. As a result, the usage of the informative signal is beneficial for the architect.

The provided expression demonstrates the role of ρ_H and ρ_L, which effectively capture the relative importance that each renovation type assigns to the design discount. If ρ_H is too high relatively to ρ_L, then the incentive cost outweighs the benefit of information provision. In particular, consider the case of perfect positive correlation $(\rho_H, \rho_L) = (1, 0)$. We can show that the interior designer's posterior when the architect does not reveal the buyer as a high renovation type remains 1/3. Hence the discount is still achieved with probability 1/2 for the high type, and 1 for the low type. Diametrically to the previous example the net impact on the architect's revenue is negative as $\rho_L - \rho_H/4 = -1/4$. Hence she is better off when not using the informative signal. Indeed, we show that this holds for any disclosure policy, that is for any distribution of signals.

The above discussion implicitly assumes that the buyer's renovation and design types are positively correlated, but what happens if $\rho_L > \rho_H$? Again information provision could be beneficial, for example consider the case of perfect negative correlation $(\rho_H, \rho_L) = (0, 1)$. However, suppose that the architect instead of revealing the high renovation type with probability $g_H = 1/2$, reveals the low one with probability $g_L = 1/2$. Hence when the architect does not reveal the buyer as a low renovation type the designer's posterior on facing a high design type is 1/3. Thus the discount is achieved with probability 1 for the high renovation type, and 1/2 for the low one. The high renovation type does not care at all about the signal, since his valuation of the interior designer's services is always low. On the other hand, the low renovation type always has a high valuation for the designer's work. As a result, the possibility of obtaining the discount allows the architect to increase the price charge to him by $(\theta_H - \theta_L)/2$. Therefore, the architect is better off when using the informative signal.

To sum up the above three examples hint that information provision is decreasing in correlation. Under perfect positive correlation the cost of incentive compatibility, of the high renovation type that is, outweighs the potential benefit, which is selling the discount to the low type. However, under imperfect positive correlation is was shown that the benefit dominates and information provision becomes optimal. Finally, under perfect negative correlation the incentive cost disappears and again information provision is optimal, however the nature of the signal that achieves that changes. The subsequent analysis will demonstrate that to some extent this intuition is relevant even when the architect is able to optimally design a mechanism to transmit information to the interior designer.

3 Model

Consider a two period model $t \in \{1, 2\}$, in which a buyer trades sequentially with sellers S_1 and S_2. Each seller supplies an indivisible good. In period 1 S_1 makes a take-it-or-leave-it offer to the buyer, while in period 2 it is S_2 that makes an offer. The sellers and the buyer are risk neutral, and all outside options are normalised to zero. Let p_t denote the price charged by S_t, and q_t the corresponding probability of supplying the good. Then the buyer's payoff from trading in period t is

$$\theta_t \, q_t - p_t$$

where $\theta_t \in \{\theta_L, \theta_H\}$, $\theta_H > \theta_L > 0$, is the value the buyer assigns to each seller's product, and the public prior on θ_1 is $\mu_0 = \Pr(\theta_1 = \theta_H)$. We allow the buyer's type θ_t to be imperfectly correlated across sellers

$$\begin{aligned} \rho_H &= \Pr(\theta_2 = \theta_H \,|\, \theta_1 = \theta_H) \\ \rho_L &= \Pr(\theta_2 = \theta_H \,|\, \theta_1 = \theta_L\,) \end{aligned}$$

Crucially, even the buyer himself will not know θ_2 before contracting with the second seller. Therefore, the model allows for some sequential resolution of uncertainty on the buyer's stochastic preferences, as in the literature of Dynamic Mechanism Design. To simplify the exposition we will initially assume that $\rho_H \geq \rho_L$, that is a period 1 high buyer type is more likely to remain so in period 2, than a low type to become one. Nevertheless, the diametrically opposite case is also considered in a separated subsection. For simplicity S_2 is only allowed to make an offer to the buyer if he accepted S_1's offer, however the analysis does not rely on this restriction.

The interaction between S_1 and the buyer will be private. Hence his report in S_1's mechanism and the outcome of the trade will not be directly observable by S_2. Nevertheless, S_1 will be able to credible convey additional information to S_2 by committing ex ante to a signal $s \in S$ with distribution $g(s \,|\, \widehat{\theta}_1)$, which is conditioned on the buyer's reported type $\widehat{\theta}_1$. It is easy to argue that the revelation principle applies in this setting. Hence in period 1 S_1 offers to the buyer a mechanism

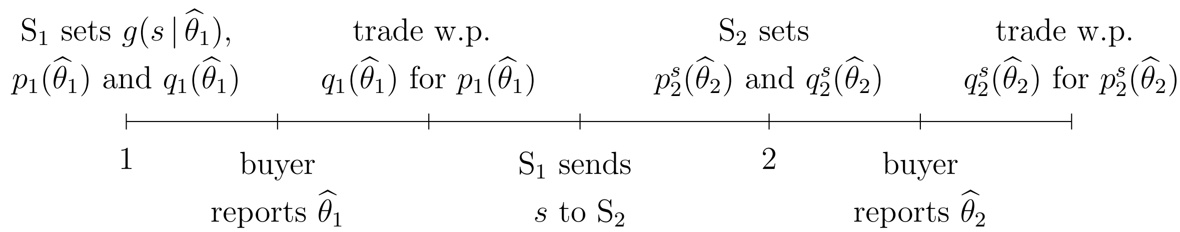
$$\left\{ \, p_1(\widehat{\theta}_1), \, q_1(\widehat{\theta}_1), \, g(s \,|\, \widehat{\theta}_1) \, \right\},$$

which specifies the price and probability of trade, and the signal's conditional distribution, respectively. Therefore the mechanism design problem of S_1 includes an information design aspect. To be more precise, S_1 is engaging in Bayesian Persuasion with S_2 about the agent's report $\widehat{\theta}_1$. In period 2, and after having received signal s, S_2 makes an offer

$$\left\{ p_2(\widehat{\theta}_2, s),\ q_2(\widehat{\theta}_2, s) \right\},$$

which depends on the signal realisation s, because this affects S_2's posterior belief on θ_1 and as a result on θ_2.

To sum up, the timing of the model is as follows. At the beginning of period 1, S_1 publicly commits to a distribution $g(s \,|\, \widehat{\theta}_1)$, and offers a corresponding mechanism to the buyer. This includes $g(s \,|\, \widehat{\theta}_1)$, as well as a choice over quantities and prices. The buyer reports $\widehat{\theta}_1$ and trades with S_1. Then the public signal s is realised and observed by S_2. At the beginning of period 2, S_2 makes a new offer to the buyer. Subsequently, the buyer reports $\widehat{\theta}_2$ and trades with S_2, at which point the game ends.



4 Analysis

Section 4.1 solves S_2's payoff maximisation problem and derives the buyer's payoff from his second trade. Subsequently, those are used to describe S_1's information provision problem. To facilitate the exposition, and because it is an interesting question on its own right, section 4.2 derives the solution of this problem under the assumption that the buyer's type is directly observable by S_1. This is equivalent to the first best solution of S_1's payoff maximisation problem. Section 4.3 reverts to the original setup, where the buyer's type is his private information, and compares its solution to the first best. Finally, 4.4 considers the case of negative correlation.

4.1 The buyer's post contractual payoff

We solve S_2's payoff maximisation problem, and derive the buyer's expected payoff from his trade with her. Let $\mu_1^s = \Pr(\theta_1 = \theta_H \,|\, s)$ denote S_2's posterior belief on the buyer's initial type after receiving s, and $\mu_2^s = \Pr(\theta_2 = \theta_H \,|\, s)$ the corresponding posterior on θ_2. Those two are connected according to

$$\mu_2^s \;=\; \mu_1^s \, \rho_H + (1 - \mu_1^s) \, \rho_L$$

S_2's problem is quite standard and a more detailed treatment can be found in the appendix. Essentially, this can be reduced to a decision between setting a high price $\overline{p} = \theta_H$, or a low price $\underline{p} = \theta_L$. In the first case only the high type buys her product, while in the second both. When her posterior on θ_2, μ_2^s, is relatively high she opts for the high price, otherwise for the low one. The cutoff in which she is indifferent between the two pricing policies is the ratio θ_L/θ_H. It will be convenient to express this in terms of the realisations of the period 1 posterior μ_1^s. Hence define

$$\mu^* \;=\; \frac{\theta_L/\theta_H - \rho_L}{\rho_H - \rho_L}$$

and note that under positive correlation[4] $\mu^* \in [0, 1]$ if and only if $\rho_L \leq \theta_L/\theta_H \leq \rho_H$. The following lemma characterises the buyer's payoff, which is the only result needed to proceed with S_1's payoff maximisation problem.

Lemma 1. *The payoff of a low buyer type under S_2 is equal to zero, while that of the high one equals*

$$Q(\mu_1^s) \;=\; \begin{cases} \theta_H - \theta_L & , \textit{if} \;\; \mu_1^s \leq \mu^* \\ 0 & , \textit{if} \;\; \mu_1^s > \mu^* \end{cases} \tag{1}$$

Proof. In Appendix A. □

The payoff of a low buyer type in period 2 is always equal to zero, as he captures no rents. Conversely, the high type's payoff is positive, but only if the posterior μ_1^s is low enough for S_2 to opt to serve both types.

Hereafter, the buyer's expected continuation payoff at the end of period 1 will be referred to as his *post contractual* payoff. It follows from the above lemma that for a period 1 low buyer type, that reported $\widehat{\theta}_1$ in S_1's mechanism, this is equal to $\rho_L \mathbb{E}_g\big[Q(\mu_1^s) \,|\, \widehat{\theta}_1\big]$, while for a high type this is $\rho_H \mathbb{E}_g\big[Q(\mu_1^s) \,|\, \widehat{\theta}_1\big]$.

[4] In the subsection where we consider negative correlation the corresponding statement is: $\mu^* \in [0, 1]$ if and only if $\rho_H \leq \theta_L/\theta_H \leq \rho_L$.

4.2 S_1's first best contract

Next, we solve S_1's payoff maximisation problem under the assumption that if the buyer opts to participate in S_1's contract, then his type is automatically reveal to her, but not to S_2. This is essentially S_1's first best contract. Hence the analysis of this subsection considers the potential benefit of information provision, without the associated incentive cost. Nonetheless, this is not only a theoretical exercise. For example, an insurance firm could ask a potential client to undergo a health examination. Therefore, in this subsection we will only consider the buyer's individual rationality constrains. Hence S_1 solves

$$\begin{aligned} \max_{p_L, p_H, q_L, q_H, g} \quad & \mu_0 p_H + (1-\mu_0)\, p_L \\ \text{s.t. } (\mathrm{IR}_L) \quad & \theta_L\, q_L - p_L + \rho_L\, \mathbb{E}_g\left[Q(\mu_1^s) \,|\, \theta_L\right] \geq 0 \\ (\mathrm{IR}_H) \quad & \theta_H q_H - p_H + \rho_H \mathbb{E}_g\left[Q(\mu_1^s) \,|\, \theta_H\right] \geq 0 \end{aligned} \qquad (\mathcal{P}_f)$$

where we write $\{p_1(\theta_L), p_1(\theta_H), q_1(\theta_L), q_1(\theta_H)\}$ as $\{p_L, p_H, q_L, q_H\}$, in order to maintain a compact notation. Both of the individual rationality constrains need to bind, as otherwise S_1 could increase p_L or p_H. Hence we can use the binding (IR_L) and (IR_H) to substitute the prices in S_1's objective function and obtain the unconstrained problem

$$\max_{q_L, q_H, g} \left\{ \begin{array}{c} \mu_0\, \theta_H\, q_H \;+\; (1-\mu_0)\, \theta_L\, q_L \\ +\mu_0 \rho_H\, \mathbb{E}_g\big[Q(\mu_1^s) \,|\, \theta_H\big] + (1-\mu_0)\rho_L \mathbb{E}_g\big[Q(\mu_1^s) \,|\, \theta_L\big] \end{array} \right\} \qquad (\mathcal{P}_f')$$

The first line of the above objective function is quite standard. It represents the surplus generated from the trade of period 1. This is optimised by supplying both types with probability one. The second line represents the additional surplus that S_1 captures from the buyer by controlling the flow of information to S_2, as well as access to her.

On the rest of this subsection we focus on S_1's information provision problem in the first best, of which the choice variable is the distribution $g(s \,|\, \theta_1)$ and the objective function the second line of $(\mathcal{P}_f')$.

$$\max_g \left\{ \mu_0 \rho_H\, \mathbb{E}_g\big[Q(\mu_1^s) \,|\, \theta_H\big] + (1-\mu_0)\rho_L \mathbb{E}_g\big[Q(\mu_1^s) \,|\, \theta_L\big] \right\} \qquad (\mathcal{G}_f)$$

The solution approach follows closely Gentzkow and Kamenica (2011). First, we rewrite the objective function of $(\mathcal{G}_f)$ as an expectation that uses the unconditional distribution of signal s. This is the ex ante probability of signal s to be realised and it is not conditioned on θ_1. Abusing notation we will denote this by

$$g(s) \;=\; \mu_0\, g(s\,|\,\theta_H) + (1-\mu_0)\, g(s\,|\,\theta_L)\,.$$

Second, we argue that instead of using the conditional distribution over signal realisations $g(s\,|\,\theta_1)$ as choice variables, S_1 can equivalently use the unconditional one over posteriors $\tilde{g}(\mu)$ with the addition of one constrain. Third, a graphical argument based on the concave closure of the reformulated objective function is used to derive the optimal $\tilde{g}(\mu)$. Finally, we identify a conditional distribution over signals that implements the optimal unconditional one over posteriors.

Lemma 2. *In the first best, S_1's information provision problem equivalently becomes*

$$\max_{g} \;\; \mathbb{E}_g[J_f(\mu_1^s)] \tag{$\mathcal{G}_f$}$$

where its point-wise value $J_f(\mu_1^s)$ is

$$J_f(\mu_1^s) \;\;=\;\; Q(\mu_1^s)\cdot[\,\mu_1^s\cdot(\rho_H-\rho_L)+\rho_L\,] \tag{2}$$

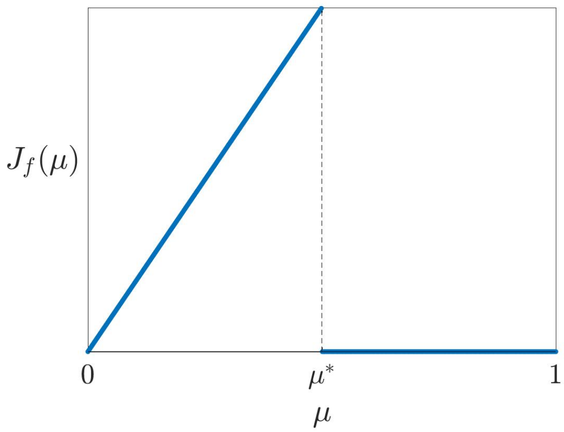
Proof. In Appendix A. □

GK show that any Bayes plausible distribution over posteriors $\tilde{g}(\mu)$ can be expressed as a conditional distribution over signals $g(s\,|\,\theta_1)$ and visa versa. A distribution over posteriors $\tilde{g}(\mu)$ is called Bayes plausible if the expected value of the posterior μ is equal to the prior μ_0. Then S_1's information provision problem equivalently becomes

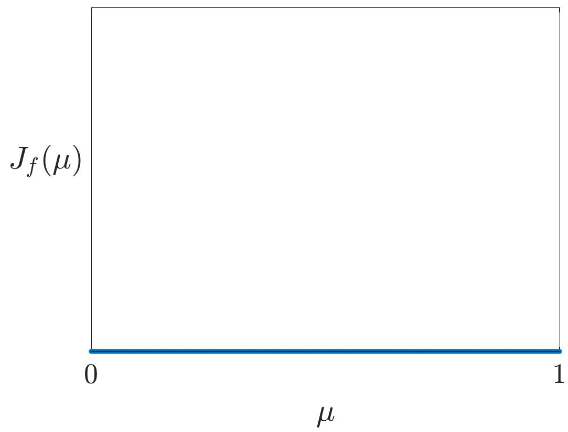
$$\max_{\tilde{g}} \;\mathbb{E}_{\tilde{g}}\big[J_f(\mu)\big] \quad \text{s.t.} \;\; \mathbb{E}_{\tilde{g}}\,[\,\mu\,] \;=\; \mu_0 \tag{$\mathcal{G}_f'$}$$

To solve this it is important to characterise the graph of J_f, which has three possible cases. Figures (1a) and (1b) demonstrate two of them. In particular, Figure (1a) is relevant when $\rho_L \leq \theta_L/\theta_H < \rho_H$, which as argued in the previous section is the only case where information provision can have an effect on prices.

To make this more clear consider the alternative shown in Figure (1b), for which $\theta_L/\theta_H < \rho_L \leq \rho_H$ has been assumed. In this case even if the buyer is revealed as a low period 1 type the probability of him to be a high type in the second period ρ_L is sufficiently high for S_2 to charge $\overline{p} = \theta_H$. As a result, the buyer's post contractual payoff is zero, irrespectively of the realisation of μ. Diametrically, if $\rho_L \leq \rho_H \leq \theta_L/\theta_H$ then even if the buyer is revealed as a high period 1 type, S_1 will still charge the low price. Hence again information provision will



(a) $\rho_L < \theta_L/\theta_H$



(b) $\rho_L \geq \theta_L/\theta_H$

Figure 1: Two cases of the graph of J_f. The inequalities $\rho_L < (\geq)\, \theta_L/\theta_H$ are equivalent to $0 < (\geq)\, \mu^*$.

have no impact on the optimal pricing strategy of S_2. Therefore a sufficient and necessary condition for information provision to have any impact is that

$$\rho_L \;\leq\; \frac{\theta_L}{\theta_H} \;<\; \rho_H \tag{3}$$

But this is only a necessary condition for information provision to strictly dominate non-disclosure. To fully solve ($\mathcal{G}'_f$) an optimality argument based on the concave closure of $J(\mu)$ will be used. This will be denoted by $\mathcal{J}_f(\mu)$ and defined as

$$\mathcal{J}_f(\mu) \;=\; \sup\big\{z \,|\, (\mu, z) \in \text{co}(J_f)\big\},$$

where $\text{co}(J_f)$ denotes the convex hull of the graph of J_f. Thus for a given $\mu \in [0,1]$, $\mathcal{J}_f(\mu)$ is the highest value that can be achieved on the vertical line that passes through μ by using any linear combination of points that are below the graph of J_f. This implies that $\mathcal{J}_f(\mu) \geq J_f(\mu)$, however the inequality could be strict. Figure (2) plots $\mathcal{J}_f$ as a dashed line whenever it is strictly bigger than J_f.

Next we want to explain how $\mathcal{J}_f$ looks like and how it is derived. Assume throughout this discussion that (3) holds, so that information provision can have an impact on prices. First we argue that if $\mu \leq \mu^*$, then there is not a linear combination of points of J_f that achieve something above $J_f(\mu)$. Since J_f is stepwise linear it suffices to only consider two posterior realisations $\mu^- \leq \mu \leq \mu^+$. If $\mu^* < \mu^+$, then the linear combination of $J_f(\mu^-)$ and $J_f(\mu^+)$ will always be below $J_f(\mu)$. If instead $\mu^+ \leq \mu^*$, then the linear combination will be equal to $J_f(\mu)$. Therefore it has to be that $\mathcal{J}_f(\mu) = J_f(\mu)$. Suppose instead that $\mu > \mu^*$, then any

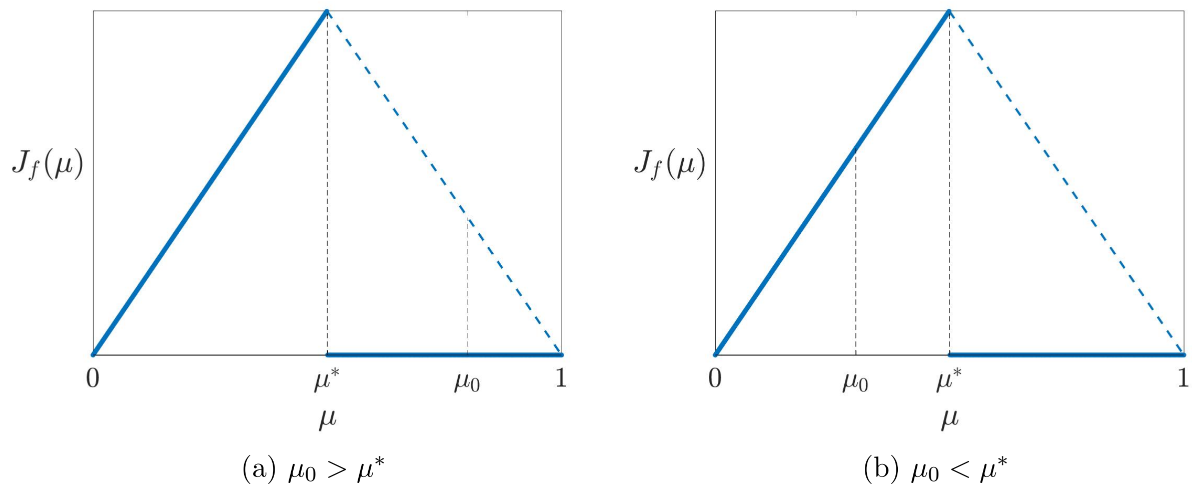


Figure 2: Two cases of the graph of J_f for which information provision can have an impact on price setting.

choice of μ^- and μ^+ such that $\mu^- \leq \mu^* < \mu < \mu^+$ will provide a linear combination higher than $J_f(\mu) = 0$. But it is always optimal to increase both μ^- and μ^+ to their maximum values, which are μ^* and 1, respectively. Therefore

$$\mathcal{J}_f(\mu) = \begin{cases} J_f(\mu) & , \text{ if } \mu \leq \mu^* \\ J_f(\mu^*) \cdot \frac{\mu - \mu^*}{1-\mu^*} & , \text{ if } \mu \geq \mu^* \end{cases}$$

The optimality argument that we use to solve S_1's information provision problem is based on the following observation: $\tilde{g}(\mu)$ solves ($\mathcal{G}'_f$) if and only if $\mathbb{E}_{\tilde{g}}[J_f(\mu)] = \mathcal{J}_f(\mu_0)$. This follows from noting that $\mathcal{J}_f(\mu_0)$ represents the maximum value that can be achieved on the vertical axis passing through μ_0 by taking linear combinations of J_f. But by its definition $\tilde{g}(\mu)$ is a set of linear combinations of J_f, and the Bayes plausibility constrain requires that those will give a value on the same vertical axis. Therefore, we can conclude that an informative signal does strictly better than no information provision if and only if $\mathcal{J}_f(\mu_0) > J_f(\mu_0)$, which is the case when $\mu_0 > \mu^*$. For a graphical illustration of this also refer to Figure (2). Before providing the main result of this subsection, we introduce the following definition.

Definition 1. An informative signal distribution $g(s \mid \theta_1)$ *strictly solves* S_1's payoff maximisation problem if it is part of one of its solutions (p_L, p_H, q_L, q_H, g), and there is no solution that uses no information provision.

Our aim is to characterise the set of parameters which can support an informative signal that strictly dominates no information provision. Using the above definition we can rule out cases where even though information provision is optimal, it is also inconsequential. An example of such a case would be when $\rho_L > \theta_L/\theta_H$, because even if S_1 reveals the buyer as a low period 1 type S_2 will still offer the high price.

Proposition 1. *In the first best, an informative signal strictly solves S_1's payoff maximisation problem ($\mathcal{P}_f$) iff*

$$\rho_L \; < \; \frac{\theta_L}{\theta_H} \; < \; \rho_H \mu_0 + \rho_L (1-\mu_0) \tag{4}$$

If this holds, then an optimal signal is $s \in \{\underline{s}, \overline{s}\}$ with distribution

$$g_f\left(\underline{s} \,|\, \theta_L\right) \;=\; 1\,, \qquad \textit{and} \qquad g_f\left(\underline{s} \,|\, \theta_H\right) \;=\; \frac{1-\mu_0}{\mu_0} \, \frac{\frac{\theta_L}{\theta_H} - \rho_L}{\rho_H - \frac{\theta_L}{\theta_H}} \tag{5}$$

In addition, it is optimal to supply both types with probability one.

Proof. In Appendix A. □

In the appendix we combine the necessary and sufficient condition for information provision to have impact on S_2's pricing policy (3) together with $\mu_0 > \mu^*$ to obtain (4). Subsequently, we show that g_f implements a randomisation between posteriors μ^* and 1, which we argued above that is the optimal distribution of posteriors when (4) holds.

When the optimal signal is informative it has a straightforward interpretation, which is that with probability $g_f\left(\overline{s} \,|\, \theta_H\right)$ S_1 reveals the high period 1 type, and otherwise says nothing. Essentially, S_1 is attempting to convince S_2 to some times offer a discount to the buyer and subsequently charges the buyer for the expected benefit of this discount. To do this she creates two signals in one of which she pools some period 1 high types with the corresponding low types. For such an informative signal to be beneficial two conditions are required. First the probability of a low period 1 type to become high period 2 type has to be low enough for S_2 to be persuaded to some times offer this discount. Second, it has to be that in the absence of persuasion S_2 would charge the high price, as otherwise the buyer would already be getting the best price possible and S_1 would have no reason to interfere.

An interesting implication of the above analysis is that information provision can be optimal even when the buyer's type is perfectly correlated across the two sellers, that is $\rho_H = 1 - \rho_L = 1$, which will be shown to not be true in S_1's second best contract. In the next section we will demonstrate that the reason why this is possible only in the first best is that the benefit of information provision can be obtained without the associated cost that the incentive compatibility constrain of the high type creates.

4.3 S_1's second best contract

Next we analyse the second best, where θ_1 is the buyer's private information. We will demonstrate that S_1's information provision problem can be manipulated in a way that allows us to use the same solution method as in the first best. S_1 solves

$$
\begin{aligned}
&\max_{p_1(\widehat{\theta}_1),\, q_1(\widehat{\theta}_1),\, g(s\,|\,\widehat{\theta}_1)} \left\{ \mu_0\, p_1(\widehat{\theta}_H) + (1-\mu_0)\, p_1(\widehat{\theta}_L) \right\} \quad \text{s.t. } (\text{IR}_L\,),\ (\text{IR}_H)\,,\\
&(\text{IC}_L\,)\ \ \theta_L\, q_1(\widehat{\theta}_L) + \rho_L\, \mathbb{E}_g\left[Q(\mu_1^s)\,|\,\widehat{\theta}_L\right] - p_1(\widehat{\theta}_L\,)\\
&\qquad\qquad \geq\ \theta_L\, q_1(\widehat{\theta}_H) + \rho_L\, \mathbb{E}_g\left[Q(\mu_1^s)\,|\,\widehat{\theta}_H\right] - p_1(\widehat{\theta}_H)\\
&(\text{IC}_H)\ \ \theta_H q_1(\widehat{\theta}_H) + \rho_H \mathbb{E}_g\left[Q(\mu_1^s)\,|\,\widehat{\theta}_H\right] - p_1(\widehat{\theta}_H)\\
&\qquad\qquad \geq\ \theta_H\, q_1(\widehat{\theta}_L) + \rho_H \mathbb{E}_g\left[Q(\mu_1^s)\,|\,\widehat{\theta}_L\,\right] - p_1(\widehat{\theta}_L)
\end{aligned}
\qquad (\mathcal{P})
$$

where the individual rationality constraints, (IR_L) and (IR_H), are as in the previous subsection. It is important to underline that the buyer's report $\widehat{\theta}_1$ affects not only the probability and price of trade in period 1, but also the distribution of the signal s. Therefore the two events on which the expectations above are conditioned are $\widehat{\theta}_1 = \widehat{\theta}_L$ and $\widehat{\theta}_1 = \widehat{\theta}_H$, which we have shortened to the reported type. On the other hand, the probabilities of becoming a period 2 high type, ρ_L and ρ_H, are only a function of the buyer's period 1 type and remain the same on both sides of the above inequalities.

To maintain a compact notation we will hereafter use θ_L and θ_H to also denote the reports $\widehat{\theta}_L$ and $\widehat{\theta}_H$, respectively, and write $\{p_L, p_H, q_L, q_H\}$ instead of $\{p_1(\widehat{\theta}_L), p_1(\widehat{\theta}_H), q_1(\widehat{\theta}_L), q_1(\widehat{\theta}_H)\}$. Similarly to the first best, it is convenient to reduce the number of constrains by substituting the transfers p_L and p_H.

Lemma 3. *In the second best, S_1's payoff maximisation problem equivalently becomes*

$$
\max_{q_L, q_H, g} \left\{ \begin{array}{c} \mu_0\, \theta_H q_H\ + (\theta_L - \mu_0\theta_H)\, q_L \\ + \mu_0\rho_H\, \mathbb{E}_g\big[Q(\mu_1^s)\,|\,\theta_H\big] + (1-\mu_0)\rho_L \mathbb{E}_g\big[Q(\mu_1^s)\,|\,\theta_L\big] \\ -\, \mu_0(\rho_H - \rho_L)\mathbb{E}_g\big[Q(\mu_1^s)\,|\,\theta_L\big] \end{array} \right\} \qquad (\mathcal{P}')
$$

$$
s.t.\quad (\theta_H - \theta_L)\,(q_H - q_L)\ \geq\ (\rho_H - \rho_L)\Big(\mathbb{E}_g\big[Q(\mu_1^s)\,|\,\theta_L\big] - \mathbb{E}_g\big[Q(\mu_1^s)\,|\,\theta_H\big]\Big) \qquad (\mathcal{P}_c)
$$

Proof. In Appendix A. □

The proof invokes the standard arguments used in mechanism design with binary type space. First, we argue that (IR_L) and (IC_H) have to bind. Second, we use the two equations to obtain transfers $\{p_L, p_H\}$ as functions of polices $\{q_L, q_H, g\}$. Substituting those in S_1's objective function and in (IC_L) gives the objective function of ($\mathcal{P}'$) and its constrain ($\mathcal{P}_c$), respectively.

S_1's objective function is quite similar with the first best, however its value is reduced because of the rents captured by the period 1 high type. Those consists of two parts, the first of which is generated from the trade of S_1's product. This appears on the first line of the objective function of ($\mathcal{P}'$) reducing the marginal benefit from trading with the low type q_L. The second part of the buyer's rents in period 1 are due to his post contractual payoff. To be more specific, those are generated from the fact that when $\rho_H > \rho_L$ the value that each type assigns to the possibility of obtaining a discount from S_2 is different, and part of what S_1 sells is this discount.

It follows from ($\mathcal{P}'$) that it is always optimal to supply the high type with probability one, as increasing q_H not only increases the objective function of ($\mathcal{P}'$) but also loosens ($\mathcal{P}_c$). The same is not true for the probability of supplying the low type q_L. If ($\mathcal{P}_c$) was ignored, then the point-wise optimal q_L would be

$$\bar{q}_L = \begin{cases} 1 & , \text{if} \quad \mu_0 \leq \theta_L/\theta_H \\ 0 & , \text{if} \quad \mu_0 \geq \theta_L/\theta_H \end{cases} \tag{6}$$

However, we will shortly demonstrate that the above will not always be implementable together with an informative signal.

In the rest of the analysis we derive the point-wise optimal signal distribution $g(s \mid \theta_1)$. In other words we solve S_1's information provision problem while ignoring the constrain ($\mathcal{P}_c$). The proof of the proposition of this subsection, which can be found in the appendix, shows that whenever ($\mathcal{P}_c$) binds S_1 prefers to decrease q_L instead of altering the point-wise optimal signal distribution $g(s \mid \theta_1)$. Define the information provision problem of S_1 as

$$\max_g \left\{ \begin{array}{c} \rho_L \cdot \mathbb{E}_g\big[Q(\mu_1^s) \mid \theta_L\big] \\ -\mu_0 \rho_H \cdot \Big(\mathbb{E}_g\big[Q(\mu_1^s) \mid \theta_L\big] - \mathbb{E}_g\big[Q(\mu_1^s) \mid \theta_H\big]\Big) \end{array} \right\} \tag{$\mathcal{G}$}$$

the objective function of which follows from gathering terms on the last two lines of the objective function of ($\mathcal{P}'$).

To better understand the incentives of S_1 to provide information remember that in the first best this is done in order to create a discount for the buyer. Suppose then that in the absence of information provision S_2 opts for the high price. Then the buyer's rents from the second period are zero, henceforth the objective function of ($\mathcal{G}$) would be zero. Is it possible for S_1 to do better? The answer depends on the relative size of ρ_L and ρ_H. An informative signal could achieve strictly positive $\mathbb{E}_g\big[Q(\mu_1^s) \mid \theta_L\big]$ and $\mathbb{E}_g\big[Q(\mu_1^s) \mid \theta_H\big]$, but it is not obvious what impact this would have on S_1's payoff.

Remark 1. For any choice of $g(s \,|\, \theta_1)$:

$$\mathbb{E}_g\big[Q(\mu_1^s) \,|\, \theta_H\big] \;\leq\; \mathbb{E}_g\big[Q(\mu_1^s) \,|\, \theta_L\big] \tag{7}$$

and the inequality is strict if $g(s \,|\, \theta_1)$ generates an impact on S_2's price.

Proof. It is without loss of generality to consider a signal $s_1, ..., s_n$ that induces progressively higher posteriors $\mu_1^{s_1} < ... < \mu_1^{s_n}$. But this implies strict monotone likelihood ratio dominance

$$\frac{g(s_n \,|\, \theta_H)}{g(s_n \,|\, \theta_L)} \;>\; \frac{g(s_{n'} \,|\, \theta_H)}{g(s_{n'} \,|\, \theta_L)} \quad \Leftrightarrow \quad n \;>\; n' \,,$$

which in turn implies that the CDF conditioned on θ_H strictly first order stochastically dominates the one conditioned on θ_L :

$$G(s \,|\, \theta_H) \;<\; G(s \,|\, \theta_L) \quad \text{for all} \quad s \;<\; s_n$$

But $Q(\mu_1^s)$ is non-increasing, which implies (7). To show that it holds with strict inequality when the signal generates an impact on prices, note that in this case $G(s \,|\, \theta_L)$ will have strictly more mass than $G(s \,|\, \theta_H)$ on the posteriors below μ^*. □

The above remark brings to the front the tension between the benefit of information provision and its potential cost. Its benefit is that it some times persuades S_2 to provide a discount to the buyer. In the first best, S_1 could capture all the expected benefit of this discount. On the other hand, in the second best not only is she not able to capture the expected benefit of the high period 1 type, but she actually has to leave some additional rents to him in order to not pretend to be a low one. This is because any informative signal will induce a better distribution for the low type, since S_2 is persuaded to offer the discount exactly when she assigns a higher probability of facing a buyer whose valuation of her product is low. We can use the above discussion to derive the equivalent of Theorem 1 of Calzolari and Pavan (2006) within our framework.

Remark 2. Suppose that the buyer's type is perfectly correlated across sellers, that is $\rho_L = 0$ and $\rho_H = 1$, then no information provision is optimal.

Proof. Follows trivially from noting that under perfect correlation ($\mathcal{G}$) becomes

$$\max_g \quad -\mu_0 \cdot \Big(\mathbb{E}_g\left[Q(\mu_1^s) \,|\, \theta_L\right] \,-\, \mathbb{E}_g\left[Q(\mu_1^s) \,|\, \theta_H\right]\Big)$$

the objective function of which is always non-positive, as argued in Remark 1. But no information provision gives always at least zero, thus it is optimal. □

Similarly to CP we showed that under perfect positive correlation privacy is optimal. As argued above when $\rho_L = 0$ the potential benefit of information provision becomes zero, while the associated incentive cost is the highest possible. Hence when we restrict ourselves to the region of positive correlation, some degree of stochasticity on the buyer's preferences is a necessary condition for information provision to be optimal.

We move on to deriving the solution of S_1's information provision problem ($\mathcal{G}$) under any imperfect positive correlation, that is under the restriction that $\rho_L \leq \rho_H$. As in the first best, to solve ($\mathcal{G}$) we start by rewriting its objective function as an expectation that uses the unconditional distribution $g(s)$.

Lemma 4. *In the second best, S_1's information provision problem equivalently becomes*

$$\max_{g} \ \mathbb{E}_g[J(\mu_1^s)] \tag{$\mathcal{G}$}$$

where its point-wise value $J(\mu_1^s)$ is

$$J(\mu_1^s) \ = \ \frac{\rho_H - \rho_L}{1-\mu_0} \cdot Q(\mu_1^s) \cdot \left(\mu_1^s - \frac{\mu_0 \rho_H - \rho_L}{\rho_H - \rho_L}\right) \tag{8}$$

Proof. In Appendix A. □

S_1's point-wise post contractual payoff J has two components. The first represents the surplus she apprehends from both types through managing their access to S_2 and transmitting information to her. The second represents the rents captured by the period 1 high type. The latter results on $J_f(\mu_1^s) > J(\mu_1^s)$ on the region of posteriors which achieve a discount. Crucially, for those posteriors $J(\mu)$ could even be negative. Under no information provision, i.e. when $\Pr(\mu_1^s = \mu_0) = 1$, S_1's post contractual payoff is equal to $\rho_L \cdot Q(\mu_0)$. In this case, S_1 essentially charges both period 1 types the post contractual payoff of the low one. Therefore this is her benefit from selling access to S_2. However, as we argued in the previous subsection whenever

$$\mu_0 \, \rho_H + (1-\mu_0) \, \rho_H \ > \ \frac{\theta_L}{\theta_H}$$

holds, this implies $Q(\mu_0) = 0$ because S_2 will charge $\overline{p} = \theta_H$ for her good. Thus, the buyer captures no surplus from the second period and the same is true for S_1. As argued in the initial example, depending on ρ_L and ρ_H it is possible for S_1 to do better by some times creating a discount, which is achieved by providing an informative signal to S_2. Effectively,

S_1 engages in Bayesian persuasion with S_2, and the underline state variable is the buyer's period 1 reported type, which he is incentivised to truthfully report in S_1's mechanism.

Similarly to before, we express S_1's information provision problem as a choice of distributions over posteriors $\tilde{g}(\mu)$. Hence it equivalently becomes

$$\max_{\tilde{g}} \ \mathbb{E}_{\tilde{g}}\big[J(\mu)\big] \quad \text{s.t.} \quad \mathbb{E}_{\tilde{g}}[\,\mu\,] = \mu_0 \tag{$\mathcal{G}'$}$$

We solve this by invoking the optimality condition $\mathbb{E}_{\tilde{g}}\big[J(\mu)\big] = \mathcal{J}(\mu_0)$, where

$$\mathcal{J}(\mu) \ = \ \sup\big\{z \,|\, (\mu, z) \in \text{co}(J)\big\}$$

denotes the concave closure of J. To obtain the functional form of $\mathcal{J}$, the shape of J needs to be characterised. Throughout the following analysis maintain the supposition that $\rho_L \leq \theta_L/\theta_H < \rho_H$, so that information provision has an impact on prices. Then the graph of J has two possible cases depending on the relative size of $\mu_0\rho_H$ and θ_L/θ_H. First assume that $\mu_0\rho_H < \theta_L/\theta_H$, in which case J looks very similar to J_f. A representative graph is given in Figure (3a). J is linear and increasing for posteriors in $[0, \mu^*]$, strictly positive at μ^*, and equals zero for posterios in $(\mu^*, \mu]$. This means that the analysis of the first best extends to this subcase of the second best. In particular, the concave closure of J is

$$\mathcal{J}(\mu) \ = \ \begin{cases} J(\mu) & , \text{if} \quad \mu \leq \mu^* \\ J(\mu^*) \cdot \frac{\mu - \mu^*}{1 - \mu^*} & , \text{if} \quad \mu \geq \mu^* \end{cases}$$

from which we infer that information provision is strictly optimal when

$$\rho_L \ < \ \frac{\theta_L}{\theta_H} \ < \rho_H\, \mu_0 + \rho_L\,(1 - \mu_0)$$

On the other hand, if $\mu_0\rho_H \geq \theta_L/\theta_H$ then J reaches μ^* while still being non-positive, as shown in Figure (3b). Hence, it never becomes strictly positive. Simple algebra gives that in this case $\mu_0 \geq \mu^*$, hence no information provision is optimal. Intuitively, S_1 can always achieve at least zero under non-disclosure, hence there is no benefit from creating an informative signal.

The connection between the first and second best solution of S_1's information provision problem is very similar to the corresponding ones of its optimal supply problem. In the first best of the latter S_1 supplies her product to both types with probability one, whereas in its second best the low types gets the product only if its virtual type $\theta_L - \mu_0\theta_H$ is positive. Interestingly, in S_1's information provision problem a 'virtual type' also appears. This is

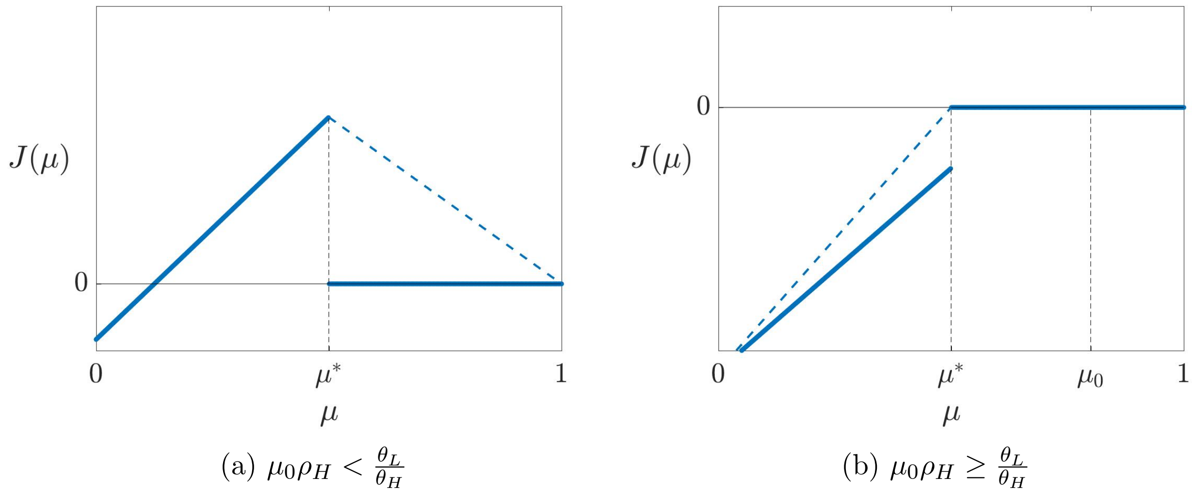


Figure 3: A representative graph of J. The dashed line denotes its concave closure, when this is above J. For the second graph we can show that $\mu_0 \geq \mu^*$ always holds.

because in ($\mathcal{G}$), $\mathbb{E}_g[Q(\mu_1^s)\,|\,\theta_L]$ is multiplied by $\rho_L - \mu_0\rho_H$, whereas in ($\mathcal{G}_f$) by $(1-\mu_0)\rho_L$. However, in this case information could be supplied even if this 'virtual type' was negative. This is because the decision to transmit information cannot be taken independently for each type. Hence the additional incentive to increase $\mu_0\rho_H\mathbb{E}_g[Q(\mu_1^s)\,|\,\theta_H]$ could skew S_1's decision towards an informative signal.

Proposition 2. *In the second best, an informative signal strictly solves S_1's payoff maximisation problem ($\mathcal{P}$) iff*

$$\max\{\rho_L, \rho_H\mu_0\} \;<\; \frac{\theta_L}{\theta_H} \;<\; \rho_H\mu_0 + \rho_L(1-\mu_0) \tag{9}$$

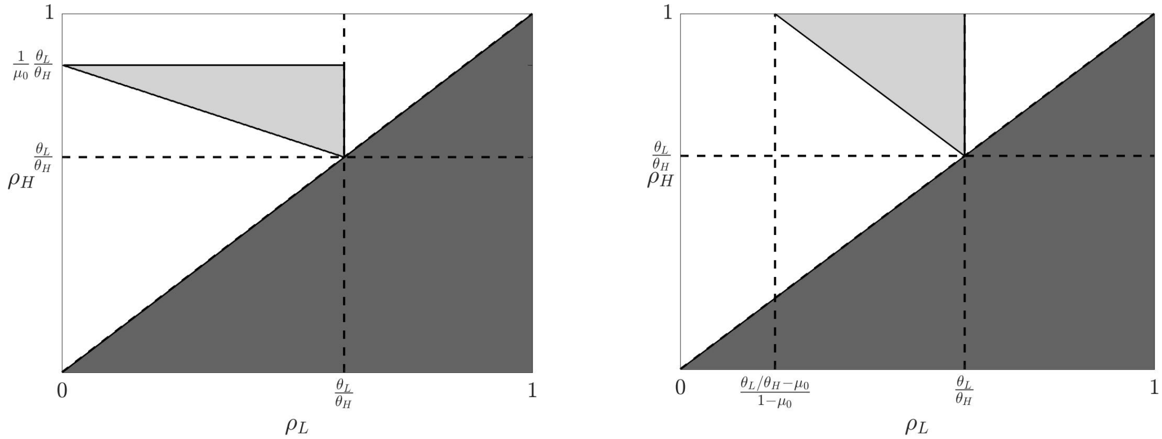
If this holds, then an optimal signal is $s \in \{\underline{s}, \overline{s}\}$ with distribution

$$g^*(\underline{s}\,|\,\theta_L) \;=\; 1 \quad \textit{and} \quad g^*(\underline{s}\,|\,\theta_H) \;=\; \frac{1-\mu_0}{\mu_0}\,\frac{\frac{\theta_L}{\theta_H}-\rho_L}{\rho_H-\frac{\theta_L}{\theta_H}} \tag{10}$$

In addition, the high type is always supplied $q_H^* = 1$. Under no information provision the low type's optimal supply schedule is the point-wise optimal (6). However, under information provision it becomes

$$q_L^* \;=\; \begin{cases} 1-(\rho_H-\rho_L)[1-g^*(\underline{s}\,|\,\theta_H)] & ,\ \textit{if}\ \ \mu_0 < \theta_L/\theta_H \\ 0 & ,\ \textit{if}\ \ \mu_0 \geq \theta_L/\theta_H \end{cases} \tag{11}$$

Proof. In Appendix A. □



(a) $\mu_0 > \theta_L/\theta_H$ (b) $\mu_0 < \theta_L/\theta_H$

Figure 4: The light grey area is the set of points for which information provision is strictly optimal in the second best. The dark grey is the area of negative correlation.

There are a few interesting observations to make. First, the necessary and sufficient condition for information provision to be strictly optimal in the second best (9) defines a convex sets of transitioning probabilities ρ_L and ρ_H. Figures (4a) and (4b) demonstrate its triangular shape. Its vertical side is the axis passing from θ_L/θ_H due to the restriction that $\rho_L < \theta_L/\theta_H$. Its diagonal is representing the second inequality of (9), which can equivalently be written as

$$\rho_H \;>\; \frac{\theta_L}{\theta_H}\frac{1}{\mu_0} - \rho_L \cdot \frac{1-\mu_0}{\mu_0} \tag{12}$$

from which we can see that the bottom corner of this triangle will always by at $(\frac{\theta_L}{\theta_H}, \frac{\theta_L}{\theta_H})$.

Second, we can note that (9) is obtained by imposing $\rho_H\mu_0 < \theta_L/\theta_H$ on the corresponding condition of the first best (4). In Figure (4a), which assumes that $\mu_0 > \theta_L/\theta_H$ this is represented by the top side of the triangle. In particular, the corresponding set of the first best would also include all the area above it. On the other hand, if $\mu_0 < \theta_L/\theta_H$, then this implies $\rho_H\mu_0 < \theta_L/\theta_H$ and as a result the set defined by the first and second best conditions is the same, as shown in Figure (4b). Therefore, in the second best information provision is strictly optimal for a smaller set of transitioning probabilities, but only if $\mu_0 > \theta_L/\theta_H$.

Third, whenever information provision is strictly optimal in the second best the optimal distribution over posteriors is the same with that of the first best. This is a randomisation between the biggest reputation that still persuades S_2 to offer the low price, which is μ^*, and one. However, this does not mean that S_1's payoff is the same, since her second best payoff will be strictly smaller.

Forth, we want to consider the welfare implications of our results. In particular, we would like to see how information provision affects the buyer's welfare. To do this we compare his payoff to that of an alternative model where information provision would not be possible. We restrict our attention to when information provision is optimal for S_1, in the original model, as otherwise we know that the buyer's payoff would be the same.

Corollary 1. *Suppose that information provision is optimal for S_1. Then the payoff of a low period 1 type is the same regardless of if information provision is possible or not. However, a high period 1 type is strictly better off when information provision is possible.*

Proof. In Appendix A. □

To further elaborate on this result note that introducing the possibility of information exchange between S_1 and S_2 leaves the payoff of a low period 1 type buyer unchanged, as his individual rationality constrain binds. However, the payoff a high type increases, since he captures additional rents. This result relies on the fact that the buyer is perfectly aware of when and how his purchase history from the first seller is shared with the second. This is exactly what legislations such as the European General Data Protection Regulation aim to achieve.

4.4 Negative correlation

Next we consider the case of negative correlation, that is we solve S_1's payoff maximisation problem in the second best ($\mathcal{P}$) under the assumption that $\rho_L > \rho_H$. In this case, the buyer's payoff from his second trade is

$$Q^-(\mu_1^s) \;=\; \mathbb{1}\{\mu_1^s \geq \mu^*\} \cdot (\theta_H - \theta_L)$$

since in contrast to the case of positive correlation S_2 is convinced to offer the low price when her posterior on the period 1 type is relatively high.

Our main difficulty is that when an informative signal is provided it is not obvious which period 1 type is really the 'high' type. This is because it could be that a period 1 low type is better off than a period 1 high type due to the former's post contractual payoff being higher than the latter's. Despite that, we can show that the representation of Lemma 3 can be extended to the case of negative correlation with the addition of one constrain.

Lemma 5. *In the second best and under negative correlation, S_1's payoff maximisation problem ($\mathcal{P}$) equivalently becomes ($\mathcal{P}'$) subject to ($\mathcal{P}_c$) and*

$$(\theta_H - \theta_L) \cdot q_L \;\geq\; (\rho_L - \rho_H) \cdot \mathbb{E}_g\big[Q^-(\mu_1^s) \,|\, \theta_L\big] \qquad (\mathcal{P}_h)$$

Proof. In Appendix A. □

The proof shows that there are two sets of potential solutions for ($\mathcal{P}$). In the first it is the (IR_L) and (IC_H) that bind, whereas in the second it is the (IR_H) and (IC_L). Nevertheless, when we restrict attention in the second set we obtain a maximum that is on the boundary of the first set. Henceforth, it is without loss to assume that the (IR_L) and (IC_H) bind. This is the premise on which the proof of Lemma 3 relies, which is why its result extends here. However, in this setting it is not necessarily true that the individual rationality constrain of the high type will be satisfied, hence we need to introduce the constrain ($\mathcal{P}_h$).

Proposition 3. *In the second best and under negative correlation, an informative signal strictly solves* S_1*'s payoff maximisation problem* ($\mathcal{P}$) *iff*

$$\rho_H \; < \; \frac{\theta_L}{\theta_H} \; < \; \rho_H \mu_0 + \rho_L(1-\mu_0) \tag{13}$$

If this holds, then an optimal signal is $s \in \{\underline{s}, \overline{s}\}$ *with distribution*

$$g^-(\underline{s}\,|\,\theta_H) \;=\; 1 \quad \textit{and} \quad g^-(\underline{s}\,|\,\theta_L) \;=\; \frac{\mu_0}{1-\mu_0}\,\frac{\frac{\theta_L}{\theta_H}-\rho_H}{\rho_L - \frac{\theta_L}{\theta_H}} \tag{14}$$

Moreover, the above signal also strictly solves S_1*'s first best payoff maximisation problem* ($\mathcal{P}_f$) *under negative correlation.*

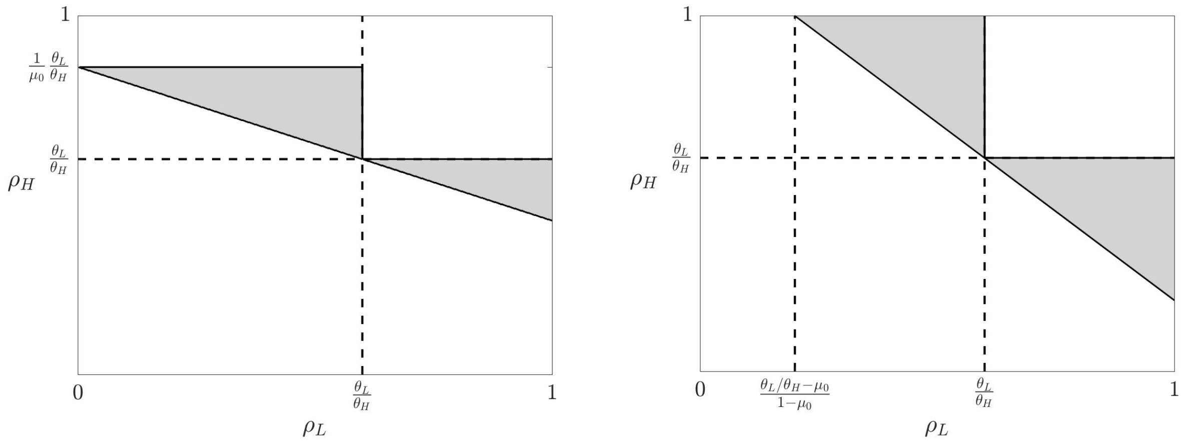
In addition, the high type is always supplied $q_H^- = 1$*. Under no information provision the low type's optimal supply schedule is the point-wise optimal* (6)*. However, under information provision it becomes*

$$q_L^- \;=\; \begin{cases} 1-(\rho_L-\rho_H)[1-g^-(\underline{s}\,|\,\theta_L)] & , \textit{ if } \;\mu_0 < \theta_L/\theta_H \\ (\rho_L-\rho_H)g^-(\underline{s}\,|\,\theta_L) & , \textit{ if } \;\mu_0 \geq \theta_L/\theta_H \end{cases}$$

Proof. In Appendix A. □

First, note that under negative correlation the informative signal implements a randomisation between posteriors 0 and μ^*. This is the opposite of that of positive correlation, which was randomising between μ^* and 1. The reason is that under negative correlation S_2 is persuaded to offer the discount for high realisations of μ_1^s instead of the low ones, because a high period 1 type is less likely to have a high valuation for her product.

Second, under negative correlation the same informative signal strictly solves both the first and second best. This is because we showed in Lemma 5 that S_1 still solves ($\mathcal{P}'$) subject to ($\mathcal{P}_c$). But we know from the previous subsection that a necessary and sufficient condition



(a) $\mu_0 > \theta_L/\theta_H$ (b) $\mu_0 < \theta_L/\theta_H$

Figure 5: The light grey area is the set of points for which information provision is optimal. Both positive and negative correlation are considered.

for the incentive cost to be low enough for information provision to be beneficial for S_1 is that $\rho_H\mu_0 < \theta_L/\theta_H$. But this is implied by the necessary condition for information provision to have an impact on prices, which is

$$\rho_H \ < \ \frac{\theta_L}{\theta_H} \ < \ \rho_L \tag{15}$$

Figures (5a) and (5b) provide the full set of transitioning probabilities for which information provision is optimal. These are obtained by extending the diagonals of the north-west triangles, which continue to represent (12), and bounding the remain area from above with the θ_L/θ_H horizontal axis, which represents the $\rho_H < \theta_L/\theta_H$ restriction. As explain before, the south-east triangles are the same for both the first and second best payoff maximisation problems.

Notably, if the slope of the diagonal was negative enough, then the south-east corner of the box would be included in the shared area. This point represents the case of perfect negative correlation, which was also covered in Proposition 2 from Calzolari and Pavan (2006). In the setup considered here, the equivalent of this result follows as a corollary of Proposition 3.

Remark 3. Suppose that the buyer's type is perfectly negatively correlated across sellers, that is $\rho_L = 1$ and $\rho_H = 0$. Then an informative signal strictly solves S_1's payoff maximisation problem ($\mathcal{P}$) iff

$$\frac{\theta_L}{\theta_H} \ < \ 1 - \mu_0 \tag{16}$$

Diametrically, to the case of perfect positive correlation, which was considered in Remark 2, here we obtain that an informative signal could be optimal even if the buyer's preferences were non-stochastic. This is because under perfect negative correlation the period 1 high type assigns zero value to the second trade, since his second period type will always be low. As a result, S_1 can create an informative signal while paying zero extra rents. Hence, S_1's information provision problem in the second best is identical to that of the first best. But in the first best an informative signal could be optimal even if the buyer's preferences were non-stochastic, which explains the above result.

Finally, we want to consider the effect of information provision on the buyer's welfare under negative correlation. As in the previous subsection, we compare the buyer's welfare in our model to an alternative one where information provision would not be possible.

Corollary 2. *Suppose that information provision is optimal for S_1. Then the payoff of a low period 1 type is the same regardless of if information provision is possible or not. The same is true for a high type if $\mu_0 \geq \theta_L/\theta_H$. However if $\mu_0 < \theta_L/\theta_H$, then a high type is strictly worse off when information provision is possible.*

Proof. In Appendix A. □

Similarly to the previous subsection, the payoff of the low period 1 type is always zero, since his individual rationality constrain binds. The crucial difference, compared to the case of positive correlation, is that the high type is worse off when information provision is possible. This is because there is a disagreement between the two types on the relative value they assign to the trade of the first and second period. The high period 1 type prioritises the first period, whereas the low the second. Therefore, when using an informative signal S_1 can exploit this misalignment to reduce the rents captured by the high type. Despite that, S_1 is only exploiting this benefit of information provision when an informative signal is already optimal in the first best. That is reducing the high type's rents is not the primary reason she discloses information, but she still enjoys it as an indirect effect. Finally, we should point out that S_1 does not achieve her first best payoff, since the low type is not served with probability one.

5 Extensions

5.1 Selling information

Throughout the previous section we assumed that S_1 could only indirectly profit from providing information to S_2. That is by charging the buyer for the discounts the informative signal could generate. However, information disclosure also generates an expected benefit for S_2, since she can use this to adjust her price and sell more frequently. Therefore, S_1 could capture part of this benefit by charging S_2 for the signal she provides. Let $\gamma \in [0,1]$ denote the proportion that S_1 captures from the expected benefit that information provision creates for S_2. Thus, setting $\gamma = 0$ would collapse the extended model to the baseline one, which allocates all the bargaining power to S_2. Diametrically, setting $\gamma = 1$ would give all the bargaining power to S_1, who would capture from S_2 all the expected benefit of her signal.

Proposition 4. *The solution of the extended model, where S_1 can directly profit from selling information to S_2, is identical to that of the baseline one under both positive and negative correlation and for all $\gamma \in [0,1]$.*

Proof. In Appendix B. □

The proof demonstrates that even though the existence of direct selling motives is beneficial for S_1, and it affects her incentives to provide information, this effect is not strong enough to alter the distribution of her optimal signal. The approach that we use in the proof is identical to that of the previous section. Henceforth, our characterisation of the subsets of priors and transitioning probabilities for which an informative signal is strictly optimal, and the distribution of this signal, is robust to direct selling motives.

Hence the welfare implications of our baseline model are still relevant. Interestingly, our analysis hints that regulations aiming at restricting information exchange between sellers should not focus on banning or reducing monetary transactions. This is because allowing the first seller to directly benefit from selling information does not affect her optimal signal. More importantly, we have already demonstrated that information exchange is not necessarily adverse for the consumers.

5.2 Static imperfect correlation

The main analysis assumes that the buyer's type evolves dynamically between the two offers. Hence the buyer only learns his second period type when trading with S_2. However, we could imagine an alternative model where the buyer would knew both his types in the first period. Hence in this case the buyer's type would be static, but it would have two imperfectly correlated components. This would effectively result on a four element type space $(\theta_1, \theta_2) \in \big\{(\theta_L, \theta_L), (\theta_H, \theta_L), (\theta_L, \theta_H), (\theta_H, \theta_H)\big\}$. We focus mainly on the case that we deemed the most reasonable by assuming that

$$\max\left\{\rho_L \, , \, \frac{\theta_L}{\theta_H}\right\} \; < \; \rho_H \tag{17}$$

which incorporates not only positive correlation, but also that if a buyer is revealed as a high period 1 type, then this suffices for the second seller to charge the high price.

Proposition 5. *Suppose that the buyer knows both his types when trading with the first seller, and that (17) holds. Then no information provision is optimal. However, if (17) does not hold then there are parameters for which information provision becomes strictly optimal.*

Proof. In Appendix B. □

Our proof proceeds as follows. First, we argue in a way similar to Remark 1 that it is not possible for the seller to transmit information on the buyer's second type. To understand this note that if this was the case the signal distribution of (θ_H, θ_L) would first order stochastically dominate that of (θ_H, θ_H). Hence the latter's second period payoff is higher under the former's distribution. In addition, the first period preferences of (θ_H, θ_H) are fully aligned with those of (θ_H, θ_L), since their valuation of the first good is the same. Hence S_1 does not have a tool at her disposal to separate them when providing information on the second period type. The same is true for (θ_L, θ_H) and (θ_L, θ_L)

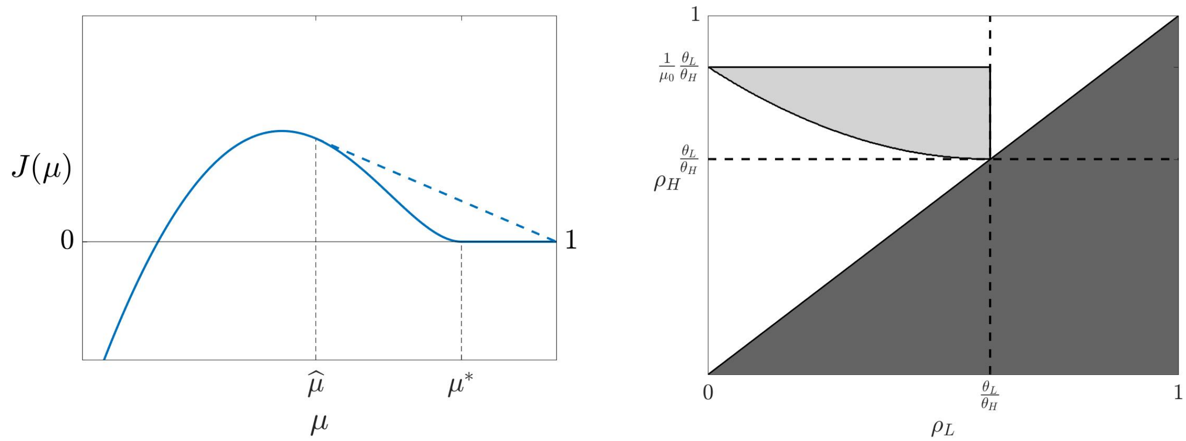
Second, suppose that S_1 tries to include all types in her contract. Then she cannot charge (θ_L, θ_H) extra for the possibility of a discount, since he could pretend to be (θ_L, θ_L) who assigns zero value to this. But similarly to the baseline model selling this discount to the low type is the only potential benefit of information provision, hence non-disclosure is optimal. There is an alternative way for S_1 to potentially increase her payoff, which is to exclude (θ_L, θ_L) from her contract. However, in that case it is only the high period 1 type that may assign the low valuation to the second seller's good. But because of (17) even if the first seller reveals the high period 1 type, this will still not be enough to generate a discount. Hence, we conclude that no information provision is optimal.

However, information provision may be optimal if (17) does not hold and the prior on the period 1 type is high enough for the first seller to not want to supply it. This is because in that case revealing a high type generates a discount, and the cost of excluding (θ_L, θ_L) from the contract is zero.

5.3 Isoelastic cost and other extensions

In an alternative extension we allow for each seller to supply a continuous quantity of a good[5] under an isoelastic cost function. We restrict attention to the case of positive correlation, for which we obtain similar results to the baseline model. The analysis is similar to that

[5] We can equivalently interpreted this as producing a single good and choosing its quality level.



(a) J in the 2nd best

(b) Area of information provision

Figure 6: The above plots consider the case where q_t is produced under an isoelastic cost. On the right plot, the light grey area is the set of points for which information provision is optimal. Only positive correlation is considered.

of the previous section, but due to its size it has been moved to Appendix C. Figure (6a) gives the the graph of J in the second best, which is the equivalent of Figure (3a) from the baseline model. In this cases the concave closure is built by finding the tangency point $\widehat{\mu}$. The set of transitioning probabilities for which information provision is optimal is as shown in (6b), which closely resembles (4a).

On the accompanying online appendix we further extend the above model by allowing the interaction of the first seller with the buyer to span over multiply periods. Since the buyer's payoff evolves stochastically during his contract under S_1, this extension is similar to the model of Battaglini (2005). However, the aforementioned paper does not include a second seller, and as a result the corresponding information provision problem does not exist. We show that S_1's preferred signal is informative for almost any choice of correlations across periods and sellers, with the noteworthy exemption of perfect positive correlation.

In a different section of the online appendix we show that the above multi-period model can also incorporated moral hazard. Therefore, it can be used to describe labour contracts. In this case, the information exchange is between employers and on the inferred ability of an employee. In a yet different section we provide some preliminary results on endogenous termination times and how the incentives of the employer to terminate her contract with the employee are affected by the fact that the second employer can use this termination time as a signal. A similar setup has been considered in Garrett and Pavan (2012), but instead of the employee being offered a new contract they assume a constant continuation value.

In the final section of the online appendix we allow for the buyer's type to be continuous.

We will not be able to derive the optimal signal for a generic distribution on the buyer's valuation and its evolution, however we will extend Calzolari and Pavan (2006) by providing some sufficient conditions for no disclosure to be optimal even under stochastic preferences.

6 Conclusions

In this paper we considered the following question: "Should a seller disclose private data on her clients?" We argued that selectively disclosing some data has two benefits. First, the seller can charge other firms for this information. Second, she can persuade those other firms to offer discounts to her customers, which she can subsequently use to increase her own prices. However, information disclosure entails an incentive cost, since it skews the choices of the seller's customers towards cheaper options. Our main contribution is to show that in a natural economic setting in which the buyer is uncertain about his future valuations information disclosure is optimal for a substantial set of environments.

A Proofs of section 4

Proof of Lemma 1. The revelation principle applies, hence it is without loss to focus on direct and incentive compatible mechanisms. To make the notation more compact, write the reported type as a subscript. S_2's revenue maximisation problem is the following one

$$\begin{aligned}
\max_{p_L, p_H, q_L, q_H} \quad & \mu_2^s\, p_H + (1-\mu_2^s)\, p_L \\
\text{s.t. } (\text{IR}_L) \quad & \theta_L q_L - p_L \geq 0 \\
(\text{IR}_H) \quad & \theta_H q_H - p_H \geq 0 \\
(\text{IC}_L) \quad & \theta_L q_L - p_L \geq \theta_L q_H - p_H \\
(\text{IC}_H) \quad & \theta_H q_H - p_H \geq \theta_H q_L - p_L
\end{aligned}$$

Assuming that (IR_L) does not bind leads to a contradiction. Subsequently, this can be used to show that (IC_H) has to bind. Hence the above simplifies to the unconstrained maximisation problem

$$\max_{q_L, q_H} \; \mu_2^s\, \theta_H\, q_H + (\theta_L - \mu_2^s \theta_H)\, q_L$$

As a result the unique solution is to set $q_H^* = 1$, while the optimal probability of supplying the low type is

$$q_L^* = \begin{cases} 1 & \text{, if } \mu_2^s \leq \theta_L/\theta_H \\ 0 & \text{, if } \mu_2^s \geq \theta_L/\theta_H \end{cases}$$

This is implementable, because substituting the above solutions in (IC_L) gives

$$\begin{aligned} \theta_L q_L - p_L \geq \theta_L q_H - p_H \;\Leftrightarrow\; 0 \geq \theta_L q_H - p_H \\ \Leftrightarrow\; \theta_H(q_H - q_L) + \theta_L q_L \geq \theta_L q_H \;\Leftrightarrow\; q_H \geq q_L, \end{aligned}$$

which is satisfied. Because the (IR_L) binds the low type's payoff is zero. The high type's payoff can be obtained using the (IR_L) and (IC_H) constraints, which give that

$$\theta_H q_H - p_H \overset{\text{IC}_H}{=} \theta_H q_L - p_L \overset{\text{IR}_L}{=} (\theta_H - \theta_L) q_L$$

□

Proof of Lemma 2. For any countable S, and $s \in S$ let the ex ante probability of s to be realised be denoted by

$$g(s) = \mu_0\, g(s \,|\, \theta_H) + (1 - \mu_0)\, g(s \,|\, \theta_L)\,.$$

Then for all $s \in S$ such that $g(s) \neq 0$ it follows that

$$g(s \,|\, \theta_H) = \frac{\mu_1^s}{\mu_0}\, g(s) \qquad \text{and} \qquad g(s \,|\, \theta_L) = \frac{1 - \mu_1^s}{1 - \mu_0}\, g(s)$$

$$\begin{aligned} &\mu_0 \rho_H\, \mathbb{E}_s\left[Q(\mu_1^s) \,|\, \theta_H\right] + (1 - \mu_0)\rho_L\, \mathbb{E}_s\left[Q(\mu_1^s) \,|\, \theta_L\right] \\ &\quad = \sum_{s \in S} Q(\mu_1^s) \Big\{ \mu_0 \rho_H g(s \,|\, \theta_H) + (1 - \mu_0)\rho_L g(s \,|\, \theta_L) \Big\} \\ &\quad\quad = \sum_{s \in S} Q(\mu_1^s) \Big\{ \rho_H \mu_1^s + \rho_L (1 - \mu_1^s) \Big\}\, g(s) = \mathbb{E}_g\big[J_f(\mu_1^s)\big]. \end{aligned}$$

□

Proof of Proposition 1. It follows from the discussion on the main part of the paper that information provision has an impact iff $\rho_L \leq \theta_L/\theta_H < \rho_H$, in which case the concave closure of $J_f(\mu)$ is given by

$$\mathcal{J}_f(\mu) = \begin{cases} J_f(\mu) & \text{, for } \mu \leq \mu^* \\ J_f(\mu^*)\frac{\mu - \mu^*}{1 - \mu^*} & \text{, for } \mu \geq \mu^* \end{cases} \qquad \text{where} \quad \mu^* = \frac{\theta_L/\theta_H - \rho_L}{\rho_H - \rho_L}$$

If $\mu_0 \leq \mu^*$, then $J_f(\mu_0) = \mathcal{J}_f(\mu_0)$ and as a result setting $\Pr(\mu = \mu_0) = 1$ is optimal, which is achieved when no information is provided to S_2. If $\mu_0 > \mu^*$, then the optimal $\tilde{g}(\mu)$ randomises between posteriors μ^* and 1. This is the solution of $(\mathcal{G}_f')$, where the choice variable is a distribution over posteriors $\tilde{g}(\mu)$. Hence a distribution over signals $g(s\,|\,\theta_1)$ solves $(\mathcal{G}_f)$ if and only if it results in a randomisation between posteriors μ^* and 1.

It is without loss to focus on a binary signal $s \in \{\underline{s}, \overline{s}\}$. Let $\overline{s}$ be the signal realisation that gives $\mu_1^s = 1$. Then it has to be that $g_f\,(\overline{s}\,|\,\theta_L) = 0$. Therefore, $g_f\,(\underline{s}\,|\,\theta_L) = 1$. Finally, to find $g_f\,(\underline{s}\,|\,\theta_H)$ note that this has to satisfy

$$\mu^* \;=\; \frac{\mu_0\, g_f\,(\underline{s}\,|\,\theta_H)}{\mu_0\, g_f\,(\underline{s}\,|\,\theta_H) + 1 - \mu_0}$$

This signal is a solution of $(\mathcal{G}_f)$, hence it is also part of the solution of $(\mathcal{P}_f)$.

In addition, $\mu_0 > \mu^*$ equivalently implies that $\rho_H\mu_0 + (1-\mu_0)\rho_L > \theta_L/\theta_H$. But since $\rho_H \geq \rho_L$ this last inequality implies $\rho_H > \theta_L/\theta_H$. Hence $\rho_H > \theta_L/\theta_H$ can be ignored, as it is implied by (4). Finally, if $\rho_L = \mu_0$ then the informative signal collapses to no information provision, which is way the left hand side of (4) has a strict inequality. □

Proof of Lemma 3. To achieve more compact expressions we adopt the notation

$$\begin{aligned} q_L &= q_1(\theta_L) \\ q_H &= q_1(\theta_H) \end{aligned} \qquad \text{and} \qquad \begin{aligned} \mathbb{E}_L &= \mathbb{E}_g\left[Q(\mu_1^s)\,|\,\theta_L\right] \\ \mathbb{E}_H &= \mathbb{E}_g\left[Q(\mu_1^s)\,|\,\theta_H\right] \end{aligned}$$

For convenience we copy below $(\mathcal{P})$ and all four of its constrains, using the shorter notation.

$$\begin{aligned} \max_{p_L, p_H, q_L, q_H, g} \quad & \mu_0 p_H + (1-\mu_0)\, p_L \\ \text{s.t. } (\mathrm{IR}_L) \quad & \theta_L\, q_L + \rho_L\, \mathbb{E}_L - p_L \;\geq\; 0 \\ (\mathrm{IR}_H) \quad & \theta_H q_H + \rho_H \mathbb{E}_H - p_H \;\geq\; 0 \\ (\mathrm{IC}_L) \quad & \theta_L\, q_L + \rho_L\, \mathbb{E}_L - p_L \;\geq\; \theta_L q_H + \rho_L \mathbb{E}_H - p_H \\ (\mathrm{IC}_H) \quad & \theta_H q_H + \rho_H \mathbb{E}_H - p_H \;\geq\; \theta_H q_L + \rho_H \mathbb{E}_L - p_L \end{aligned} \tag{$\mathcal{P}$}$$

Using the (IC_H) we infer that whenever the (IR_L) is satisfied the same is true for the (IR_H), since we have assumed for now that $\rho_H \geq \rho_L$. Suppose that (IR_L) was not binding on the maximum of $(\mathcal{P})$. Then S_1 could increase both p_L and p_H by the same constant $\varepsilon > 0$. This would increase her payoff and for ε small enough still satisfy all the constrains, which leads to a contradiction. Thus (IR_L) has to bind on the maximum of $(\mathcal{P})$, which gives

$$p_L \;=\; \theta_L\, q_L \;+\; \rho_L\, \mathbb{E}_L$$

Next, suppose that (IC_H) was not binding on the maximum of ($\mathcal{P}$). Then S_1 could increase p_H by a small enough constant $\varepsilon > 0$, which would increase her payoff and still satisfy all the constrains. Hence we obtain again a contradiction and (IC_H) has to bind, which gives

$$p_H \;=\; \theta_H(q_H - q_L) + \rho_H(\mathbb{E}_H - \mathbb{E}_L) + p_L$$

Substitute p_H and p_L in the objective function of ($\mathcal{P}$) to obtain

$$\begin{aligned}
\mu_0 p_H + (1-\mu_0)\, p_L \;&=\; \mu_0\big[\theta_H(q_H - q_L) + \rho_H(\mathbb{E}_H - \mathbb{E}_L)\big] + p_L \\
&= \mu_0\, \theta_H q_H \; + (\theta_L - \mu_0\theta_H)\, q_L \\
&\quad + \mu_0\rho_H\, \mathbb{E}_H + (1-\mu_0)\rho_L \mathbb{E}_L \;-\; \mu_0(\rho_H - \rho_L)\mathbb{E}_L
\end{aligned}$$

which is the objective function given on the statement of this Lemma. So far we have ignored the (IC_L). This can be rewritten as

$$\begin{aligned}
p_H - p_L \;&\geq\; \theta_L(q_H - q_L) + \rho_L(\mathbb{E}_H - \mathbb{E}_L) \\
\Leftrightarrow \theta_H(q_H - q_L) + \rho_H(\mathbb{E}_H - \mathbb{E}_L) \;&\geq\; \theta_L(q_H - q_L) + \rho_L(\mathbb{E}_H - \mathbb{E}_L) \\
\Leftrightarrow (\theta_H - \theta_L)(q_H - q_L) \;&\geq\; (\rho_H - \rho_L)(\mathbb{E}_L - \mathbb{E}_H)
\end{aligned}$$

which gives the ($\mathcal{P}_c$) constrain. □

Proof of Lemma 4. For any countable S, it has already been shown in Lemma 2 that

$$\mu_0\rho_H\, \mathbb{E}_g\big[Q(\mu_1^s)\,|\,\theta_H\big] + (1-\mu_0)\rho_L \mathbb{E}_g\big[Q(\mu_1^s)\,|\,\theta_L\big] \;=\; \sum_{s\in S} Q(\mu_1^s)\,\Big\{\,\rho_H\mu_1^s + \rho_L(1-\mu_1^s)\,\Big\}\, g(s)$$

In addition, deleting irrelevant signals from S, i.e. those that occur with zero probability on path, and substituting $g(s\,|\,\theta_L) = \frac{1-\mu_1^s}{1-\mu_0}\, g(s)$ gives

$$\begin{aligned}
\mu_0(\rho_H - \rho_L)\mathbb{E}_g\left[Q(\mu_1^s)\,|\,\theta_L\right] \;&=\; \sum_{s\in S} Q(\mu_1^s)\,\mu_0\,(\rho_H - \rho_L)\, g(s\,|\,\theta_L) \\
&= \sum_{s\in S} Q(\mu_1^s)\,\mu_0\,(\rho_H - \rho_L)\,\frac{1-\mu_1^s}{1-\mu_0}\, g(s)
\end{aligned}$$

Thus, subtracting the second part above from the first gives

$$\begin{aligned}
&\sum_{s\in S} Q(\mu_1^s)\left\{\rho_H\mu_1^s + \rho_L(1-\mu_1^s) - \mu_0\left(\rho_H-\rho_L\right)\frac{1-\mu_1^s}{1-\mu_0}\right\} g(s) \\
= &\sum_{s\in S} \frac{Q(\mu_1^s)}{1-\mu_0}\left\{(\rho_H-\rho_L)\mu_1^s(1-\mu_0) + \rho_L(1-\mu_0) - \mu_0\left(\rho_H-\rho_L\right)\left(1-\mu_1^s\right)\right\} g(s) \\
= &\sum_{s\in S} \frac{Q(\mu_1^s)}{1-\mu_0}\left\{(\rho_H-\rho_L)(\mu_1^s-\mu_0) + \rho_L(1-\mu_0)\right\} g(s) \\
= &\frac{\rho_H-\rho_L}{1-\mu_0}\sum_{s\in S} Q(\mu_1^s)\left\{\mu_1^s - \mu_0 + \frac{\rho_L(1-\mu_0)}{\rho_H-\rho_L}\right\} g(s) \\
= &\frac{\rho_H-\rho_L}{1-\mu_0}\sum_{s\in S} Q(\mu_1^s)\left\{\mu_1^s - \frac{\mu_0\rho_H-\rho_L}{\rho_H-\rho_L}\right\} g(s)
\end{aligned}$$

□

Proof of Proposition 2. Proving that the proposed signal solves ($\mathcal{G}'$) follows exactly the same argumentation as in the first best, hence it is omitted. To show that it also solves ($\mathcal{P}'$) it remains to consider how it interacts with the constrain ($\mathcal{P}_c$), which can equivalently be written as

$$q_H - q_L \;\geq\; (\rho_H-\rho_L)\Big(\mathbb{E}_g\left[\mathbb{1}\left\{\mu_1^s\leq\mu^*\right\}\,|\,\theta_L\right] - \mathbb{E}_g\left[\mathbb{1}\left\{\mu_1^s\leq\mu^*\right\}\,|\,\theta_H\right]\Big) \qquad (\mathcal{P}_c')$$

It is always optimal to set $q_H^* = 1$. First suppose that $\theta_L/\theta_H \leq \mu_0$, which implies that the point-wise optimal supply for the low type is $q_L = 0$. But in this case ($\mathcal{P}_c'$) trivially holds, since its left hand side equals one, while its right hand side is always less than one. This proves that in this case the point-wise optimal signal also solves ($\mathcal{P}'$).

Hereafter, the proof only considers the diametrically opposite case $\theta_L/\theta_H > \mu_0$. Rewrite the right hand side of ($\mathcal{P}_c'$) as follows

$$\begin{aligned}
&(\rho_H-\rho_L)\Big(\mathbb{E}_g\left[\mathbb{1}\left\{\mu_1^s\leq\mu^*\right\}\,|\,\theta_L\right] - \mathbb{E}_g\left[\mathbb{1}\left\{\mu_1^s\leq\mu^*\right\}\,|\,\theta_H\right]\Big) \\
&\quad= \frac{\rho_H-\rho_L}{\theta_H-\theta_L}\sum_s Q(\mu_1^s)\Big(g(s\,|\,\theta_L) - g(s\,|\,\theta_H)\Big) \\
&\quad= \frac{\rho_H-\rho_L}{\theta_H-\theta_L}\sum_s Q(\mu_1^s)\left(\frac{1-\mu_1^s}{1-\mu_0} - \frac{\mu_1^s}{\mu_0}\right) g(s) \\
&\quad= \frac{\rho_H-\rho_L}{\theta_H-\theta_L}\sum_s Q(\mu_1^s)\frac{\mu_0-\mu_1^s}{(1-\mu_0)\mu_0}\, g(s)
\end{aligned}$$

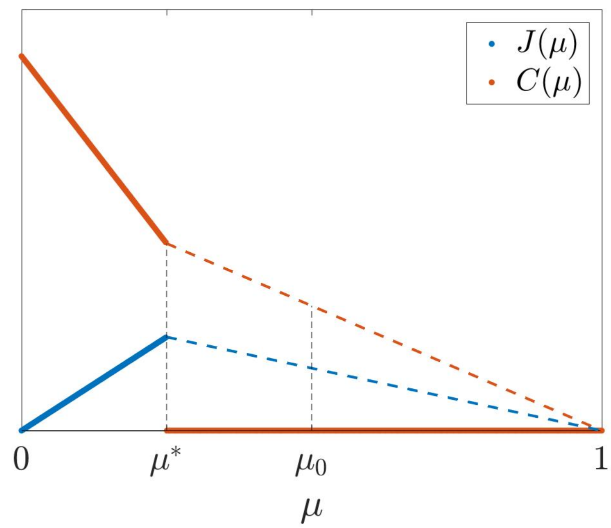


Figure 7: The point-wise post contractual payoff of S_1 $J(\mu)$, and the point-wise value of the right hand side of $(\mathcal{P}''_c)$.

As a result, the constrain becomes

$$q_H - q_L \;\geq\; \mathbb{E}_{\tilde{g}}\big[C(\mu)\big] \quad \text{where} \quad C(\mu) \;=\; \frac{\rho_H - \rho_L}{\theta_H - \theta_L}\, Q(\mu)\, \frac{\mu_0 - \mu}{(1-\mu_0)\mu_0} \qquad (\mathcal{P}''_c)$$

Under no information provision the right hand side is equal to $C(\mu_0) = 0$, therefore if no information provision solves $(\mathcal{G}')$, then it also solves $(\mathcal{P}')$.

Instead suppose that information provision solves $(\mathcal{G}')$. Similarly to the first best, this implies that $\mu_0 > \mu^*$. Therefore the point-wise optimal supply, when $\theta_L/\theta_H > \mu_0$, is $q_H = q_L = 1$, while under the point-wise optimal signal

$$\mathbb{E}_{\tilde{g}}\big[C(\mu)\big] \;=\; \tilde{g}(\mu^*)C(\mu^*) \;>\; 0$$

Hence under the point-wise optimal solutions the left hand side of $(\mathcal{P}''_c)$ would be zero, whereas the right hand side would be positive.

Note that both $J(\mu)$ and $C(\mu)$ are piecewise linear below and above μ^*. Hence it is without loss to consider a distribution that induces only two posteriors $\mu^- \leq \mu^* < \mu^+$. For this part of the proof it will potentially be useful to refer to Figure (7). Suppose that $\mu^- < \mu^*$, which implies

$$J(\mu^-) < J(\mu^*) \quad \text{and} \quad C(\mu^-) > C(\mu^*).$$

Since the expectations $\mathbb{E}_{\tilde{g}}[J(\mu)]$ and $\mathbb{E}_{\tilde{g}}[C(\mu)]$ are linear combinations of the value of each function at μ^- and zero, it follows that switching to $\mu^- = \mu^*$ is strictly better. This is because it increases $\mathbb{E}_{\tilde{g}}[J(\mu)]$ and it decreases $\mathbb{E}_{\tilde{g}}[C(\mu)]$. Then S_1 is always better off by leaving the supply of the high type and the low realisation of the buyer's posterior on their point-wise optimal values. Hence $(\mathcal{P}')$ can be written as

$$\begin{aligned} \max_{q_L,\tilde{g}^-} \quad & (\theta_L - \mu_0\theta_H)\, q_L + \tilde{g}^-\, J_1(\mu^*) \\ \text{s.t.} \quad & 1 - q_L \;\geq\; \tilde{g}^-\, C(\mu^*) \end{aligned}$$

where $\tilde{g}^- = \tilde{g}(\mu^*)$. When $\mu^- = \mu^*$, $\tilde{g}^-$ is a bijection of μ^+, hence it will be convenient to use the former as a choice variable. The objective function and the constrain are linear in both choice variables. Hence S_1 prefers to decrease q_L instead of $\tilde{g}^-$ when

$$\begin{aligned} \theta_L - \mu_0\theta_H \;\leq\; \frac{J_1(\mu^*)}{C(\mu^*)} \quad &\Leftrightarrow \quad \theta_L - \mu_0\theta_H \;\leq\; \frac{\frac{\rho_H-\rho_L}{1-\mu_0} Q(\mu^*)\left(\mu^* - \frac{\mu_0\rho_H-\rho_L}{\rho_H-\rho_L}\right)}{\frac{\rho_H-\rho_L}{\theta_H-\theta_L}\frac{Q(\mu^*)}{(1-\mu_0)\mu_0}\left(\mu_0-\mu^*\right)} \\ \Leftrightarrow \quad & \frac{\frac{\theta_L}{\theta_H}-\mu_0}{1-\frac{\theta_L}{\theta_H}}\left(1-\frac{\mu^*}{\mu_0}\right) \;\leq\; \mu^* - \frac{\mu_0\rho_H-\rho_L}{\rho_H-\rho_L} \\ \Leftrightarrow \quad \frac{\theta_L}{\theta_H} - \mu_0 \;\leq\; & \left[\left(\frac{\theta_L}{\theta_H}\frac{1}{\mu_0} - 1\right) + \left(1 - \frac{\theta_L}{\theta_H}\right)\right]\mu^* - \frac{\mu_0\rho_H-\rho_L}{\rho_H-\rho_L}\left(1-\frac{\theta_L}{\theta_H}\right) \\ \Leftrightarrow \quad & \frac{\theta_L}{\theta_H} - \mu_0 \;\leq\; \frac{\theta_L}{\theta_H}\frac{1-\mu_0}{\mu_0}\mu^* - \frac{\mu_0\rho_H-\rho_L}{\rho_H-\rho_L}\left(1-\frac{\theta_L}{\theta_H}\right) \\ \Leftrightarrow \quad & 1-\mu_0 \;\leq\; \frac{\theta_L}{\theta_H}\frac{1-\mu_0}{\mu_0}\mu^* + \frac{\rho_H(1-\mu_0)}{\rho_H-\rho_L}\left(1-\frac{\theta_L}{\theta_H}\right) \end{aligned}$$

But remember that $\mu^* = \frac{\theta_L/\theta_H - \rho_L}{\rho_H-\rho_L}$. Hence the above equivalently becomes

$$1 \;\leq\; \frac{\theta_L}{\theta_H}\frac{1}{\mu_0}\frac{\theta_L/\theta_H-\rho_L}{\rho_H-\rho_L} + \frac{\rho_H}{\rho_H-\rho_L}\left(1-\frac{\theta_L}{\theta_H}\right)$$

In addition, the above trade off is only relevant if $\frac{\theta_L}{\theta_H}\frac{1}{\mu_0} > 1$. Hence it suffices that

$$\rho_H - \rho_L \;\leq\; \frac{\theta_L}{\theta_H} - \rho_L + \rho_H - \rho_H\frac{\theta_L}{\theta_H} \quad\Leftrightarrow\quad 0 \;\leq\; \frac{\theta_L}{\theta_H}(1-\rho_H)$$

which holds. As a result the point-wise optimal signal solves $(\mathcal{P}')$, while the probability of supplying the good to the low type becomes

$$q_L^* \;=\; \begin{cases} 1 - (\rho_H - \rho_L)[1 - g^*(\underline{s}\,|\,\theta_H)] & \text{, if } \mu_0 < \theta_L/\theta_H \\ 0 & \text{, if } \mu_0 \geq \theta_L/\theta_H \end{cases}$$

Finally, any solution of ($\mathcal{P}'$) is also a solution of ($\mathcal{P}$), which completes the proof. □

Proof of Corollary 1. The payoff a low type is always zero, since the (IR_L) binds. For the high type, when information provision is possible we can use the binding (IC_H) and (IR_L) to obtain that his payoff is

$$(\theta_H - \theta_L) \cdot q_L^* + (\rho_H - \rho_L) \cdot \mathbb{E}_{g^*}[Q(\mu_1^s) \,|\, \theta_L] \tag{18}$$

where q_L^* and g^* are as given in Proposition 2. First, suppose that $\mu_0 \geq \theta_L/\theta_H$. If information provision was not possible, then the high type's payoff would be zero. However, when it is possible $q_L^* = 0$ and (18) becomes

$$(\rho_H - \rho_L) \cdot \mathbb{E}_g[Q(\mu_1^s) \,|\, \theta_L] > 0$$

hence he is strictly better off. Second, suppose that $\mu_0 < \theta_L/\theta_H$. If information provision was not possible, then the high type's payoff would be $\theta_H - \theta_L$. However, when it is possible

$$q_L^* \;=\; 1 - (\rho_H - \rho_L) \cdot [1 - g^*(\underline{s} \,|\, \theta_H)]$$

Hence (18) becomes

$$\begin{aligned}(\theta_H - \theta_L) \cdot \Big[1 - (\rho_H - \rho_L)[1 - g^*(\underline{s} \,|\, \theta_H)] + (\rho_H - \rho_L) \cdot g^*(\underline{s} \,|\, \theta_L)\Big] \\ = \; (\theta_H - \theta_L) \cdot \Big[1 + (\rho_H - \rho_L) \cdot g^*(\underline{s} \,|\, \theta_H)\Big]\end{aligned}$$

which is bigger than $\theta_H - \theta_L$, hence again he is better off. □

Proof of Lemma 5. For the convenience of the reader ($\mathcal{P}$) together with all four of its constrains is copied below

$$\begin{aligned}
\max_{p_L, p_H, q_L, q_H, g} \quad & \mu_0 p_H + (1-\mu_0)\, p_L \\
\text{s.t. } (\text{IR}_L) \quad & \theta_L\, q_L + \rho_L\, \mathbb{E}_L - p_L \;\geq\; 0 \\
(\text{IR}_H) \quad & \theta_H q_H + \rho_H \mathbb{E}_H - p_H \;\geq\; 0 \\
(\text{IC}_L) \quad & \theta_L\, q_L + \rho_L\, \mathbb{E}_L - p_L \;\geq\; \theta_L q_H + \rho_L \mathbb{E}_H - p_H \\
(\text{IC}_H) \quad & \theta_H q_H + \rho_H \mathbb{E}_H - p_H \;\geq\; \theta_H q_L + \rho_H \mathbb{E}_L - p_L
\end{aligned} \tag{$\mathcal{P}$}$$

where the compact notation introduced in the proof of Lemma 3 is used. First, we want to demonstrate that under negative correlation ($\mathcal{P}$) has two possible families of solutions.

First, suppose that

$$(\theta_H - \theta_L)q_L \;\geq\; (\rho_L - \rho_H)\mathbb{E}_L \tag{19}$$

which equivalently implies

$$\theta_H\, q_L + \rho_H \mathbb{E}_L - p_L \;\;\geq\;\; \theta_L\, q_L + \rho_L \mathbb{E}_L - p_L \;,$$

This together with the (IC$_H$) give

$$\theta_H\, q_H + \rho_H \mathbb{E}_H - p_H \;\;\geq\;\; \theta_L\, q_L + \rho_L \mathbb{E}_L - p_L$$

Next it is shown that a necessary implication of (19) is that (IR$_L$) and (IC$_H$) need to bind. Suppose (IR$_L$) does not bind, then prices $\tilde{p}_L = p_L + \varepsilon$ and $\tilde{p}_H = p_H + \varepsilon$, for $\varepsilon > 0$ small enough, increase S$_1$'s payoff and satisfy all of the above constrains, which leads to a contradiction. Suppose (IC$_H$) does not bind, then $\tilde{p}_H = p_H + \varepsilon$, again for $\varepsilon > 0$ small enough, increases S$_1$'s payoff and satisfies all of the above constrains, which leads to a contradiction. Hence rewriting the binding (IC$_H$) and the (IC$_L$) gives

$$\theta_H(q_H - q_L) + \rho_H(\mathbb{E}_H - \mathbb{E}_L) \;=\; p_H - p_L \;\geq\; \theta_L(q_H - q_L) + \rho_L(\mathbb{E}_H - \mathbb{E}_L)$$

and combining those two together gives

$$(\theta_H - \theta_L)(q_H - q_L) \;\geq\; (\rho_L - \rho_H)(\mathbb{E}_H - \mathbb{E}_L) \tag{20}$$

which is an equivalent expression of ($\mathcal{P}_c$). Finally, adding up the initial assumption (19) and (20) gives

$$(\theta_H - \theta_L)q_H \;\geq\; (\rho_L - \rho_H)\mathbb{E}_H \tag{21}$$

which will be used at the end of the proof.

Next the diametrically opposite case of solutions is considered, but it will be more convenient to start by assuming that it is (21) that holds with the opposite direction. That is suppose that the solution of ($\mathcal{P}$) satisfies

$$(\theta_H - \theta_L)q_H \;\leq\; (\rho_L - \rho_H)\mathbb{E}_H \tag{22}$$

which equivalently implies

$$\theta_L q_H + \rho_L \mathbb{E}_H - p_H \;\geq\; \theta_H q_H + \rho_H \mathbb{E}_H - p_H$$

This together with the (IC_L) give

$$\theta_L q_L + \rho_L \mathbb{E}_L - p_L \ \geq \ \theta_H q_H + \rho_H \mathbb{E}_H - p_H$$

Then following an argumentation similar to that of the previous case we can show that (IR_H) and (IC_L) need to bind. Hence rewriting the binding (IC_L) and the (IC_H) gives

$$\theta_H(q_H - q_L) + \rho_H(\mathbb{E}_H - \mathbb{E}_L) \ \geq \ p_H - p_L \ = \ \theta_L(q_H - q_L) + \rho_L(\mathbb{E}_H - \mathbb{E}_L)$$

and combining those two together gives (20) again. Combining the supposition (22) with (20) gives

$$(\theta_H - \theta_L)q_L \ \leq \ (\rho_L - \rho_H)\mathbb{E}_L\,,$$

which is the opposite of the first supposition considered (19). But this implies that the solution of ($\mathcal{P}$) satisfies

$$(\theta_H - \theta_L)q_L \ \geq \ (\rho_L - \rho_H)\mathbb{E}_L \quad \Leftrightarrow \quad (\theta_H - \theta_L)q_H \ \geq \ (\rho_L - \rho_H)\mathbb{E}_H$$

Hence, there are only two possible families of solutions for ($\mathcal{P}$). Either (19) holds, in which case (IR_L) and (IC_H) bind, or it holds with the reverse direction in which case (IR_H) and (IC_L) bind.

Next we want to show that (19) will always hold on the maximum. The proof is by contradiction. Suppose that (19) does not hold, then it has to be that it holds with the opposite direction. Therefore, the (IR_L) and (IC_H) bind. Those two give p_L and p_H, which can be substituted in the objective function of ($\mathcal{P}$) and (IC_L) to obtain

$$\max_{q_L, q_H, g} \left\{ \begin{array}{c} [\theta_H - (1-\mu_0)\theta_L]q_H \ + (1-\mu_0)\theta_L\, q_L \\ + [\rho_H - (1-\mu_0)\rho_L]\mathbb{E}_H + (1-\mu_0)\rho_L \mathbb{E}_L \end{array} \right\} \tag{$\mathcal{P}'_L$}$$

$$\text{s.t.} \quad (\theta_H - \theta_L)\,(q_H - q_L) \ \geq \ (\rho_L - \rho_H)(\mathbb{E}_H - \mathbb{E}_L) \tag{20}$$

$$(\theta_H - \theta_L)q_H \ \leq \ (\rho_L - \rho_H)\mathbb{E}_H \tag{22}$$

where (22) is used instead of the inverse of (19), since those two are equivalent.

First, note that setting $q_H = \frac{\rho_L - \rho_H}{\theta_H - \theta_L}\,\mathbb{E}_H$ is always optimal. This is because increasing q_H relaxes (22) and increases the objective function of ($\mathcal{P}'_L$). Hence (22) has to bind. Second, note that increasing q_L increases the objective function of ($\mathcal{P}_L$). Hence ($\mathcal{P}_c$) has to bind. But this implies that (19) also binds.

However, this means that the maximum under (22) is also available under the initial supposition that (19) holds. Therefore, it is without loss to assume that it is the (IR_L) and (IC_H) that bind. But in this case the representation of Lemma 3 is still relevant. However, it is not necessarily true that (IR_H) will be satisfied, hence this has to be added to the contains. But using the binding (IR_L) and (IC_H) we can see that (IR_H) equivalently becomes (19). □

Proof of Proposition 3. First, we find the point-wise optimal signal under negative correlation. S_1's information provision problem is still the same as in ($\mathcal{G}$), however its reformulated version will be different. Following exactly the same approach as in the previous subsection we can show that ($\mathcal{G}$) equivalently becomes

$$\max_{\tilde{g}} \; \mathbb{E}_{\tilde{g}}\big[J^{-}(\mu)\big] \quad \text{s.t.} \quad \mathbb{E}_{\tilde{g}}[\,\mu\,] = \mu_0 \tag{$\mathcal{G}^-$}$$

where

$$J^{-}(\mu) \;=\; \frac{\rho_H - \rho_L}{1-\mu_0}\, Q^{-}(\mu) \left(\mu - \frac{\mu_0\rho_H - \rho_L}{\rho_H - \rho_L}\right)$$

This is almost identical to the functional form of $J(\mu)$ under positive correlation, however $Q(\mu) = \mathbb{1}\{\mu \leq \mu^*\}(\theta_H - \theta_L)$ has to be replaced with

$$Q^{-}(\mu) \;=\; \mathbb{1}\{\mu \geq \mu^*\}(\theta_H - \theta_L)$$

The reason for this alteration is that when the correlation is negative S_2 offers the discount for high realisations of μ_1^s, instead of low ones. Therefore the graph of $J^{-}(\mu)$ under negative correlation can be obtained by flipping that of $J(\mu)$ around a vertical axis passing from μ^*. Hence $J^{-}(\mu)$ is flat at zero for $\mu < \mu^*$, then jumps to $J^{-}(\mu^*)$ and is subsequently decreasing. Thus a necessary condition for information provision to be strictly optimal is that

$$J^{-}(\mu^*) \,>\, 0 \quad \Leftrightarrow \quad \frac{\theta_L}{\theta_H} \,>\, \mu_0\rho_H$$

Moreover, following the same argumentation as in the case of positive correlation we get than information provision is strictly point-wise optimal if and only if

$$\max\{\rho_H, \mu_0\rho_H\} \;<\; \frac{\theta_L}{\theta_H} \;<\; \mu_0\rho_H + (1-\mu_0)\rho_L$$

But this is implied by

$$\rho_H \;<\; \frac{\theta_L}{\theta_H} \;<\; \mu_0\rho_H + (1-\mu_0)\rho_L$$

which is also a necessary and sufficient condition for information provision to be optimal in the first best under negative correlation. Hence under negative correlation the point-wise optimal signal in the first and second best is the same. The second inequality ensures that the discount is not achieved in the absence of information provision, while the first that a signal that achieves this discount can be constructed. This is a randomisation between posterior 0 and μ^*, and it is easy to check that the signal distribution provided in the statement of the proposition achieves that.

However, note that same way as in the case of positive correlation the constrain ($\mathcal{P}_c$) will bind under the point-wise optimal solutions when $\mu_0 < \theta_L/\theta_H$. Despite that, the argumentation used in the proof of Proposition 2 is still relevant and it shows that S_1 would always be better off by decreasing the probability of supplying the low type, instead of alternating the point-wise optimal signal.

On the other hand, when $\mu_0 > \theta_L/\theta_H$ and the point-wise optimal signal entails information provision, it is the ($\mathcal{P}_h$) that binds. Hence we have that

$$(\theta_H - \theta_L)q_L \;=\; (\rho_L - \rho_H)\mathbb{E}_L$$

where we use the compact notation introduced in the proof of Lemma 3. In this case, it is always optimal to set $q_H = 1$ and ($\mathcal{P}_c$) holds. Hence substitute the above equality in the objective function of ($\mathcal{P}'$) to obtain the unconstrained information provision problem

$$\max_g \;\left\{(\theta_L - \mu_0\theta_H)\cdot\frac{\rho_L - \rho_H}{\theta_H - \theta_L}\,\mathbb{E}_L + \mathbb{E}_g\left[J^-(\mu_1^s)\right]\right\} \tag{$\widetilde{\mathcal{G}}$}$$

where the second term is the standard part of the information provision problem of S_1, while the first is the one introduced by the binding ($\mathcal{P}_h$). We can further simplify this by noting that

$$\mathbb{E}_L \;=\; \mathbb{E}_g\left[Q^-(\mu_1^s)\cdot\frac{1-\mu_1^s}{1-\mu_0}\right]$$

Hence same way as before we can reformulate the above as a choice over unconditional posterior distributions $\tilde{g}(\mu)$.

$$\max_{\tilde{g}}\;\mathbb{E}_{\tilde{g}}\left[\widetilde{J}^-(\mu)\right] \quad \text{s.t.} \quad \mathbb{E}_{\tilde{g}}[\,\mu\,] = \mu_0 \tag{$\widetilde{\mathcal{G}}'$}$$

where

$$\begin{aligned}
\widetilde{J}^-(\mu) \;&=\; Q^-(\mu)\cdot(\theta_L - \mu_0\theta_H)\cdot\frac{\rho_L-\rho_H}{\theta_H-\theta_L}\cdot\frac{1-\mu}{1-\mu_0} + J^-(\mu)\\
&=\; Q^-(\mu)\cdot\frac{\rho_L-\rho_H}{1-\mu_0}\cdot\left\{\frac{\rho_L-\mu_0\rho_H}{\rho_L-\rho_H} - \mu + \frac{\mu_0\theta_H-\theta_L}{\theta_H-\theta_L}\cdot\big(\mu-1\big)\right\}\\
&=\; Q^-(\mu)\cdot\frac{\rho_L-\rho_H}{1-\mu_0}\cdot\left\{\frac{\rho_L-\mu_0\rho_H}{\rho_L-\rho_H} - \frac{\mu_0\theta_H-\theta_L}{\theta_H-\theta_L} - \mu\cdot\frac{(1-\mu_0)\cdot\theta_H}{\theta_H-\theta_L}\right\}
\end{aligned}$$

Hence, $\widetilde{J}^{-}(\mu)$ is equal to zero in $[0, \mu^*)$ and decreasing in $[\mu^*, 1]$. In addition, we can show that $\widetilde{J}^{-}(\mu^*) > 0$. It suffices to prove this for $\mu^* = 1$, in which case the inequality becomes

$$\frac{\rho_L - \mu_0 \rho_H}{\rho_L - \rho_H} \; > \; \frac{\mu_0 \theta_H - \theta_L}{\theta_H - \theta_L} + 1 \cdot \frac{(1-\mu_0)\cdot \theta_H}{\theta_H - \theta_L} \; = \; 1$$

which holds as $\mu_0 \in (0,1)$ and $\rho_L > \rho_H$. Therefore, $\widetilde{J}^{-}$ has the same shape with J^{-}, as a result they share the same optimal unconditional distribution over posteriors. Finally, in this case q_L is found be using the binding ($\mathcal{P}_h$) instead of ($\mathcal{P}_c$). □

Proof of Corollary 2. The payoff a low type is always zero, since his individual rationality constrain binds. For the high type, first suppose that $\mu_0 \geq \theta_L/\theta_H$. Thus if information provision was not possible, then his payoff would be zero. But even when information provision is possible we have argued in the proof of Proposition 3 that his individual rationality constrains binds, hence his payoff is again zero. Finally, suppose that $\mu_0 < \theta_L/\theta_H$. If information provision was not possible, then the high type's payoff would be $\theta_H - \theta_L$. However, when it is possible we can use the binding (IC_H) and (IR_L) to obtain that his payoff is

$$(\theta_H - \theta_L)\cdot q_L^{-} + (\rho_H - \rho_L)\cdot \mathbb{E}_{g^-}[Q(\mu_1^s)\,|\,\theta_L] \tag{23}$$

where q_L^{-} and g^{-} are as given in Proposition 2. In particular,

$$q_L^{-} \; = \; 1 - (\rho_L - \rho_H)\cdot[1 - g^{-}(\underline{s}\,|\,\theta_L)]$$

Hence (23) becomes

$$\begin{aligned}(\theta_H - \theta_L)\cdot \Big[1 - (\rho_L - \rho_H)\cdot[1 - g^{-}(\underline{s}\,|\,\theta_L)] + (\rho_H - \rho_L)\cdot g^{-}(\underline{s}\,|\,\theta_L)\Big] \\ = \; (\theta_H - \theta_L)\cdot \Big[1 - (\rho_L - \rho_H)\Big]\end{aligned}$$

which is smaller than $\theta_H - \theta_L$. Hence in this case the high type buyer is worse off under information provision. □

B Proofs of Section 5

B.1 Selling information

Our aim in this subsection is to prove Proposition 4. Therefore, in the subsequent analysis we allow S_1 to profit directly from selling information to S_2. Following Calzolari and Pavan

(2006) we model this with an exogenous parameter $\gamma \in [0,1]$ that denotes the part of the ex ante benefit of information provision that S_1 captures from S_2. To be more specific, S_2's payoff for given posterior μ_2^s is

$$U_2^s \;=\; \mu_2^s\,\theta_H + \max\{0\,,\,\theta_L - \mu_2^s\theta_H\}$$

In addition, let U_2^{ND} denote S_2's payoff under no disclosure. Then S_1 captures

$$\gamma\cdot\big(\mathbb{E}_g[U^s]-U^{ND}\big) \;=\; \gamma\cdot\mathbb{E}_g\left[\mathbb{1}\left(\mu_2^s \le \frac{\theta_L}{\theta_H}\right)(\theta_L-\mu_2^s\theta_H)\right] \\ -\gamma\cdot\max\Big\{0\,,\,\theta_L-\big[\mu_0\theta_H+(1-\mu_0)\theta_L\big]\theta_H\Big\}$$

We impose no restrictions on the sign of the correlation, hence to facilitate the exposition let $\underline{\rho}=\min\{\rho_L,\rho_H\}$ and $\overline{\rho}=\max\{\rho_L,\rho_H\}$. Furthermore, we adopt the compact notation introduced in the proof of Lemma 3, and define $\mathbb{1}_2(\mu_2^s) \;=\; \mathbb{1}\left(\mu_2^s \le \frac{\theta_L}{\theta_H}\right)$. Therefore, S_1 solves

$$\begin{aligned} &\max_{p_L,p_H,q_L,q_H,g} && \mu_0 p_H + (1-\mu_0)\,p_L + \gamma\cdot\mathbb{E}_g\big[\mathbb{1}_2(\mu_2^s)(\theta_L-\mu_2^s\theta_H)\big] \\ &\text{s.t. } (\mathrm{IR}_L\,) && \theta_L\,q_L \,+\, \rho_L\,\mathbb{E}_L - p_L \;\ge\; 0 \\ &\quad\;\; (\mathrm{IR}_H) && \theta_H q_H + \rho_H\mathbb{E}_H - p_H \;\ge\; 0 \\ &\quad\;\; (\mathrm{IC}_L\,) && \theta_L\,q_L + \,\rho_L\,\mathbb{E}_L\, - p_L \;\ge\; \theta_L q_H + \rho_L\mathbb{E}_H - p_H \\ &\quad\;\; (\mathrm{IC}_H) && \theta_H q_H + \rho_H\mathbb{E}_H - p_H \;\ge\; \theta_H q_L + \rho_H\mathbb{E}_L - p_L \end{aligned} \tag{Γ}$$

where the constant part of $\gamma\cdot\big(\mathbb{E}_g[U^s]-U^{ND}\big)$ is suppressed.

Lemma B.6. *(Γ) equivalently becomes*

$$\max_{q_L,q_H,g}\left\{\begin{array}{c} \mu_0\,\theta_H q_H \,+ (\theta_L-\mu_0\theta_H)\,q_L \\ +\,\mu_0\rho_H\,\big(\mathbb{E}_H-\mathbb{E}_L\big)+\rho_L\,\mathbb{E}_L+\gamma\cdot\mathbb{E}_g\big[\mathbb{1}_2(\mu_2^s)(\theta_L-\mu_2^s\theta_H)\big] \end{array}\right\} \tag{Γ'}$$

$$s.t.\quad (\theta_H-\theta_L)\,(q_H-q_L) \;\ge\; (\rho_H-\rho_L)\big(\mathbb{E}_L-\mathbb{E}_H\big) \tag{Γ_c}$$

The proof is identical to that of Lemma 5, so it is omitted. The only difference between (Γ') and the corresponding ($\mathcal{P}'$) is the addition of $\gamma\cdot\mathbb{E}_g\big[\mathbb{1}_2(\mu_2^s)(\theta_L-\mu_2^s\theta_H)\big]$ on its second line. As in the main text, we start by ignoring the constrain (Γ') and focus on finding the point-wise optimal signal, which solves

$$\max_g\;\Big\{\mu_0\rho_H\,\big(\mathbb{E}_H-\mathbb{E}_L\big)+\rho_L\,\mathbb{E}_L+\gamma\cdot\mathbb{E}_g\big[\mathbb{1}_2(\mu_2^s)(\theta_L-\mu_2^s\theta_H)\big]\Big\} \tag{$\mathcal{G}_\gamma$}$$

Lemma B.7. *S_1's information provision problem equivalently becomes*

$$\max_{g} \;\; \mathbb{E}_g[J_\gamma(\mu_2^s)] \tag{$\mathcal{G}_\gamma$}$$

where its point-wise value $J_\gamma(\mu_2^s)$ is

$$J_\gamma(\mu_2^s) \;=\; \mathbb{1}_2(\mu_2^s)\cdot\left\{\mu_2^s\theta_H\cdot\left(\frac{1-\theta_L/\theta_H}{1-\mu_0}-\gamma\right)-\theta_L\cdot\left(\frac{1-\theta_L/\theta_H}{1-\mu_0}\cdot\frac{\mu_0\rho_H}{\theta_L/\theta_H}-\gamma\right)\right\} \tag{B.24}$$

Proof. We have already shown in the proof of Lemma 4 that

$$\mu_0\rho_H\cdot\big(\mathbb{E}_H-\mathbb{E}_L\big)+\rho_L\,\mathbb{E}_L \;=\; \frac{\rho_H-\rho_L}{1-\mu_0}\cdot(\theta_H-\theta_L)\cdot\mathbb{E}_g\left[\mathbb{1}_2(\mu_2^s)\cdot\left(\mu_1^s-\frac{\mu_0\rho_H-\rho_L}{\rho_H-\rho_L}\right)\right]$$

the right hand side of which can be further manipulated to obtain

$$\mu_0\rho_H\cdot\big(\mathbb{E}_H-\mathbb{E}_L\big)+\rho_L\,\mathbb{E}_L \;=\; \frac{\theta_H-\theta_L}{1-\mu_0}\cdot\mathbb{E}_g\Big[\mathbb{1}_2(\mu_2^s)\cdot(\mu_2^s-\mu_0\rho_H)\Big]$$

Therefore, adding on the above the payoff obtained from selling information gives

$$\mathbb{E}_g\left[\mathbb{1}_2(\mu_2^s)\cdot\left\{\frac{\theta_H-\theta_L}{1-\mu_0}\cdot\Big(\mu_2^s-\mu_0\rho_H\Big)+\gamma\cdot\big(\theta_L-\mu_2^s\theta_H\big)\right\}\right] \;=$$
$$\mathbb{E}_g\left[\mathbb{1}_2(\mu_2^s)\cdot\left\{\mu_2^s\theta_H\cdot\left(\frac{1-\theta_L/\theta_H}{1-\mu_0}-\gamma\right)-\theta_L\cdot\left(\frac{1-\theta_L/\theta_H}{1-\mu_0}\cdot\frac{\mu_0\rho_H}{\theta_L/\theta_H}-\gamma\right)\right\}\right]$$

□

Notice that we conditioned the point-wise payoff J_γ on S$_2$'s posterior on θ_2 (μ_2^s) instead of that on θ_1 (μ_1^s). This will make the expressions shorter and valid for both positive and negative correlation. Next, we express ($\mathcal{G}_\gamma$) as a choice over posteriors $\tilde{g}(\mu)$ on θ_2:

$$\max_{\tilde{g}} \;\; \mathbb{E}_{\tilde{g}}\big[J_\gamma(\mu)\big] \quad \text{s.t.} \quad \begin{cases} \mathbb{E}_{\tilde{g}}[\,\mu\,]=\mu_0\rho_H+(1-\mu_0)\rho_L \\ \mu\in\big[\underline{\rho}\,,\,\overline{\rho}\big] \end{cases} \tag{$\mathcal{G}_\gamma'$}$$

where note that the posterior μ is bounded by the transitioning probabilities ρ_L and ρ_H, as it is on θ_1 instead of θ_2. For the same reason its expected value has to be equal to $\mu_0\rho_H+(1-\mu_0)\rho_L$, instead of μ_0. To solve this we invoke the optimality condition that $\mathbb{E}_{\tilde{g}}\big[J_\gamma(\mu)\big]=\mathcal{J}_\gamma\big(\mu_0\rho_H+(1-\mu_0)\rho_L\big)$, where

$$\mathcal{J}_\gamma \;=\; \sup\big\{z\,|\,(\mu,z)\in\operatorname{co}(J_\gamma)\big\},$$

denotes the concave closure of J_γ.

Next, we want to understand how the graph of J_γ looks like. It will be convenient to consider different cases on the underline parameters, each one on a corresponding lemma. Throughout, we maintain the assumption that

$$\underline{\rho} \;\leq\; \frac{\theta_L}{\theta_H} \;<\; \overline{\rho} \tag{B.25}$$

so that information provision can have an impact on prices.

Lemma B.8. *Suppose that $\rho_H\mu_0 \geq \frac{\theta_L}{\theta_H}$, then $J_2(\mu) \leq 0$ for all $\mu \in \left[\underline{\rho}\,,\,\overline{\rho}\right]$.*

Therefore, information provision can never be strictly optimal under the above parametric restriction, since no information provision gives at least zero. This is identical to our conclusion in the baseline model.

***Proof*.** First note that $\rho_H\mu_0 \geq \frac{\theta_L}{\theta_H}$ implies $\rho_H > \frac{\theta_L}{\theta_H}$, which together with (B.25) gives that $\rho_L \leq \frac{\theta_L}{\theta_H}$. Therefore, we only need to consider the case of positive correlation. As in the baseline model

$$J_\gamma(\mu) \;=\; 0\,, \quad \text{for all} \quad \mu \in \left(\frac{\theta_L}{\theta_H}\,,\,\rho_H\right]$$

Hence, it remain to prove that $J_\gamma(\mu) \leq 0$ for all $\mu \in \left[\rho_L, \frac{\theta_L}{\theta_H}\right]$. As J_γ is linear on this subset it will suffice to show that $J_\gamma\left(\frac{\theta_L}{\theta_H}\right)$ and $J_\gamma(\rho_L)$ are non-positive.

$$J_\gamma\left(\frac{\theta_L}{\theta_H}\right) \;=\; \frac{\theta_H - \theta_L}{1-\mu_0}\cdot\left(\frac{\theta_L}{\theta_H} - \mu_0\rho_H\right)$$

As a result,

$$J_\gamma\left(\frac{\theta_L}{\theta_H}\right) \;>\; 0 \quad\Leftrightarrow\quad \frac{\theta_L}{\theta_H} \;>\; \rho_H\mu_0 \tag{B.26}$$

Therefore, $J_\gamma\left(\frac{\theta_L}{\theta_H}\right) \leq 0$. Finally,

$$\begin{aligned} J_\gamma(\rho_L) \;=\; \frac{\theta_H-\theta_L}{1-\mu_0}\cdot\Big(\rho_L - \mu_0\rho_H\Big) + \gamma\cdot\big(\theta_L - \rho_L\theta_H\big) \;\leq\; 0 \\ \Leftrightarrow \quad \gamma \;\leq\; \frac{1-\theta_L/\theta_H}{1-\mu_0}\cdot\frac{\mu_0\rho_H - \rho_L}{\theta_L/\theta_H - \rho_L} \end{aligned}$$

But $\rho_H\mu_0 \geq \frac{\theta_L}{\theta_H}$ implies $\mu_0 \geq \frac{\theta_L}{\theta_H}$, as a result the right hand side of the above inequality is always greater or equal than one, from which it follows that it holds for any $\gamma \in [0,1]$. $\square$

Next we want to consider the case where $\rho_H \mu_0 < \frac{\theta_L}{\theta_H}$, which we will further break down to two subcases.

Lemma B.9. *Suppose that* $\rho_H \mu_0 < \frac{\theta_L}{\theta_H}$ *and* $\gamma < \frac{1-\theta_L/\theta_H}{1-\mu_0}$*, then*

- J_γ *is linear and increasing in* $\left[\underline{\rho}, \frac{\theta_L}{\theta_H}\right]$ *and equals zero in* $\left(\frac{\theta_L}{\theta_H}, \overline{\rho}\right]$
- *becomes strictly positive on* $\frac{\theta_L}{\theta_H}$*, that is* $J_\gamma\left(\frac{\theta_L}{\theta_H}\right) > 0$

Proof. The second bullet point follows immediately from (B.26). To prove the first bullet point note that (B.24) gives

$$J_\gamma(\mu) = \begin{cases} \mu \cdot \theta_H \cdot \left(\frac{1-\theta_L/\theta_H}{1-\mu_0} - \gamma\right) - \theta_L \cdot \left(\frac{1-\theta_L/\theta_H}{1-\mu_0} \cdot \frac{\mu_0 \rho_H}{\theta_L/\theta_H} - \gamma\right) & \text{, if } \mu \in \left[\underline{\rho}, \frac{\theta_L}{\theta_H}\right] \\ 0 & \text{, if } \mu \in \left(\frac{\theta_L}{\theta_H}, \overline{\rho}\right] \end{cases} \tag{B.27}$$

□

It will convenient to discuss the implications of the above lemma, after providing the one that characterises J_γ in the remaining parametric restriction.

Lemma B.10. *Suppose that* $\rho_H \mu_0 < \frac{\theta_L}{\theta_H}$ *and* $\gamma \geq \frac{1-\theta_L/\theta_H}{1-\mu_0}$*, then*

- J_γ *is linear, non-increasing, and strictly positive in* $\left[\underline{\rho}, \frac{\theta_L}{\theta_H}\right]$*, and equals zero in* $\left(\frac{\theta_L}{\theta_H}, \overline{\rho}\right]$
- *Extending the line that is* J_γ *in* $\left[\underline{\rho}, \frac{\theta_L}{\theta_H}\right]$ *to* $\overline{\rho}$ *would give a positive value, in other words*

$$J_\gamma\left(\frac{\theta_L}{\theta_H}\right) + J'_\gamma\left(\frac{\theta_L}{\theta_H}\right) \cdot \left(\overline{\rho} - \frac{\theta_L}{\theta_H}\right) \geq 0$$

Proof. The first bullet point follows from (B.26) and (B.27). For the second bullet point, note that the positive part of J_γ calculated at $\overline{\rho}$ becomes

$$\frac{\theta_H - \theta_L}{1-\mu_0} \cdot \left(\overline{\rho} - \mu_0 \rho_H\right) + \gamma \cdot \left(\theta_L - \overline{\rho}\,\theta_H\right) \geq 0$$

$$\Leftrightarrow \quad \frac{1-\theta_L/\theta_H}{\overline{\rho} - \theta_L/\theta_H} \cdot \frac{\overline{\rho} - \mu_0 \rho_H}{1-\mu_0} \geq \gamma \quad \Leftrightarrow \quad \frac{\overline{\rho} - \overline{\rho} \cdot \frac{\theta_L}{\theta_H}}{\overline{\rho} - \frac{\theta_L}{\theta_H}} \cdot \frac{1 - \mu_0 \cdot \frac{\rho_H}{\overline{\rho}}}{1-\mu_0} \geq \gamma$$

where both of the two fractions on the left hand side of the last inequality are greater than one, hence it has to hold. □

Therefore, in the case covered in Lemma B.9 the shape of J_γ is similar to that obtained for J in the baseline model. As a result, its concave closure will be

$$\mathcal{J}_\gamma(\mu) \;=\; \begin{cases} J_\gamma(\mu) & \text{, if} \quad \mu \leq \frac{\theta_L}{\theta_H} \\ J_\gamma(\mu^*) \cdot \frac{\mu - \frac{\theta_L}{\theta_H}}{\overline{\rho} - \frac{\theta_L}{\theta_H}} & \text{, if} \quad \mu \geq \frac{\theta_L}{\theta_H} \end{cases}$$

Interestingly, the above concave closure will also be relevant under the parametric restriction of Lemma B.10, which follows from its second bullet point.

Therefore, arguing in the same way as in the main text we get that information provision is strictly optimal if and only if

$$\max\left\{\underline{\rho}\,,\ \rho_H\mu_0\right\} \;<\; \frac{\theta_L}{\theta_H} \;<\; \rho_H\mu_0 + \rho_L(1-\mu_0)$$

in which case the optimal $\tilde{g}$ randomises between $\frac{\theta_L}{\theta_H}$ and $\overline{\rho}$.

The above discussion ignores the (Γ_c) constrain, however this is equivalent to that of the baseline case ($\mathcal{P}_c$). In addition, the incentives to maintain the point-wise optimal signal are even stronger for $\gamma > 0$. Therefore, S_1 will again opt to decrease the probability of providing the good to the low type, when necessary, instead of altering the point-wise optimal signal. Hence, the solution of this extension is identical to that of the baseline model.

B.2 Static Type

Our aim in this subsection is to prove Proposition 5. Hence we assume that the buyer knows both his types when interacting with the first seller, which effectively implies a four element space $(\theta_1, \theta_2) \in \big\{(\theta_L, \theta_L), (\theta_H, \theta_L), (\theta_L, \theta_H), (\theta_H, \theta_H)\big\}$. As in the main text, to maintain a compact notation we will write $\{p_{LL}, p_{HL}, q_{LH}, q_{HH}\}$ instead of

$$\big\{p_1(\widehat{\theta}_L, \widehat{\theta}_L), p_1(\widehat{\theta}_H, \widehat{\theta}_L), q_1(\widehat{\theta}_L, \widehat{\theta}_H), q_1(\widehat{\theta}_H, \widehat{\theta}_H)\big\}$$

Also, when it is not causing confusion we will use the non-hated types even when referring to the reports.

A trivial case in which information provision can be optimal when (17) does not hold is that of perfect negative correlation, as we have shown in Remark 3. Hence in the remaining of this subsection we will assume that (17) holds and prove that in this case no information provision is always optimal.

Lemma B.11. *Let* $\mathbb{1}_2(\mu_2^s) = \mathbb{1}\left(\mu_2^s \leq \frac{\theta_L}{\theta_H}\right)$*, then for any choice of* $g(s \,|\, \theta_1, \theta_2)$ *:*

$$\mathbb{E}_g\big[\mathbb{1}_2(\mu_2^s) \,|\, \theta_1, \theta_H\big] \;\leq\; \mathbb{E}_g\big[\mathbb{1}_2(\mu_2^s) \,|\, \theta_1, \theta_L\big] \quad \textit{for} \quad \theta_1 \in \{\theta_L, \theta_H\} \tag{B.28}$$

and the inequality is strict if $g(s \,|\, \theta_1, \theta_2)$ *generates an impact on* S_2*'s price.*

Proof. Using the towering property we can equivalently rewrite (B.28) as

$$\mathbb{E}_g\Big[\mathbb{E}_g\big[\mathbb{1}_2(\mu_2^s) \,|\, \theta_2 = \theta_H\big] - \mathbb{E}_g\big[\mathbb{1}_2(\mu_2^s) \,|\, \theta_2 = \theta_L\big] \,\Big|\, \theta_1\Big] \;\leq\; 0 \quad \text{for} \quad \theta_1 \in \{\theta_L, \theta_H\}$$

Hence it suffices to prove the statement of the lemma for the expression within the first expectation. This follows from a proof identical with that of Remark 1. □

Next, we want to consider the implication of the above lemma on the information structures that S_1 can use.

Lemma B.12. *It is without loss to focus on information structures that transmit transmits information about* θ_1*, but not about* θ_2*.*

Proof. The revelation principle holds. Hence it has to be that a (θ_H, θ_L) buyer prefers to truthfully report his type instead of (θ_H, θ_H) and visa versa

$$q_{HL}\,\theta_H - p_{HL} \;\geq\; q_{HH}\,\theta_H - p_{HH} \tag{B.29}$$

$$q_{HH}\,\theta_H + \mathbb{E}_g\left[Q(\mu_1^s) \,|\, \theta_H, \theta_H\right] - p_{HH} \;\geq\; q_{HL}\,\theta_H + \mathbb{E}_g\left[Q(\mu_1^s) \,|\, \theta_H, \theta_L\right] - p_{HL} \tag{B.30}$$

Adding those two up we get $\mathbb{E}_g\left[Q(\mu_1^s) \,|\, \theta_H, \theta_H\right] \geq \mathbb{E}_g\left[Q(\mu_1^s) \,|\, \theta_H, \theta_L\right]$. But whenever an information structure creates an impact on prices we get that the above contradicts Lemma B.11. The same contradictions can be obtained for (θ_L, θ_L) and (θ_L, θ_H).

Therefore, any informative signal that transmits information on the second period type, in a way that impacts prices, cannot be implemented. Henceforth, it is without loss to ignore such information structures. □

Thus we further simplify our notation by writing $\mathbb{E}_L$ and $\mathbb{E}_H$ to denote $\mathbb{E}_g\left[Q(\mu_1^s) \,|\, \theta_1 = \theta_L\right]$ and $\mathbb{E}_g\left[Q(\mu_1^s) \,|\, \theta_1 = \theta_H\right]$, respectively.

We start by solving S_1's problem under the assumption that she creates a contract in which all four types will participate. To facilitate the exposition we provide here the four individual rationality constrains.

$$
\begin{aligned}
(\mathrm{IR}_{LL}\ )\ \ q_{LL}\theta_L - p_{LL} &\geq 0\\
(\mathrm{IR}_{LH}\ )\ \ q_{LH}\theta_L + \mathbb{E}_H - p_{LH} &\geq 0\\
(\mathrm{IR}_{HL}\ )\ \ q_{HL}\theta_H - p_{HL} &\geq 0\\
(\mathrm{IR}_{HH})\ \ q_{HH}\theta_H + \mathbb{E}_H - p_{HH} &\geq 0
\end{aligned}
$$

To maintain a compact notation will denote the incentive compatibility constrain that ensures that (θ_H, θ_L) does not want to report (θ_L, θ_H) as IC(HL,LH). We will use the same notation to refer to the rest of the IC constrains.

Note that Lemma B.12 together with inequalities (B.29) and (B.30) give (B.31) below, while (B.32) follows from the corresponding incentive compatibility constrains of the period 1 low types.

$$
q_{HL}\,\theta_H - p_{HL} = q_{HH}\,\theta_H - p_{HH} \tag{B.31}
$$

$$
q_{LL}\,\theta_L - p_{LL} = q_{LH}\,\theta_L - p_{LH} \tag{B.32}
$$

Lemma B.13. *Suppose that S_1 offers an contract in which all four types participate. Then her payoff maximisation problem equivalently becomes*

$$
\begin{aligned}
&\max_{p,g} \Big\{\mu_0\rho_H p_{HH} + \mu_0(1-\rho_H)p_{HL} + (1-\mu_0)\rho_L p_{LH} + (1-\mu_0)(1-\rho_L)p_{LL}\Big\}\\
&s.t \quad p_{LL} = q_{LL}\theta_L\,, \quad p_{LH} = q_{LH}\theta_L\,,\\
&\qquad p_{HL} = (q_{HL} - q_{HH})\theta_H + p_{HH}\\
&\qquad (q_{HH} - q_{LL}\,)\theta_H + q_{LL}\theta_L - (\mathbb{E}_L - \mathbb{E}_H\,) \geq p_{HH} \geq q_{HH}\theta_H\\
&\qquad (q_{HH} - q_{LH})\theta_H + q_{LH}\theta_L - (\mathbb{E}_L - \mathbb{E}_H) \geq p_{HH} \geq q_{HH}\theta_H
\end{aligned}
\tag{$\mathcal{P}_4$}
$$

Proof. We first want to show that the payoff of (θ_L, θ_L) is the lowest of the four types, which will imply that (i) IR_{LL} has to bind and (ii) we can ignore the other three individual rationality constrains. This follows from the derivations below

$$
\begin{aligned}
&\text{(B.32)} \qquad q_{LH}\,\theta_L + \mathbb{E}_L - p_{LH} \geq q_{LL}\,\theta_L - p_{LL}\\
&\text{IC(HH,LL)} \quad q_{HH}\,\theta_H + \mathbb{E}_H - p_{HH} \geq q_{LL}\,\theta_H + \mathbb{E}_L - p_{LL}\\
&\qquad\qquad\qquad \Rightarrow \quad q_{HH}\,\theta_H + \mathbb{E}_H - p_{HH} \geq q_{LL}\,\theta_L - p_{LL}\\
&\text{IC(HL,LL)} \quad q_{HL}\,\theta_H - p_{HL} \geq q_{LL}\,\theta_H - p_{LL}\\
&\qquad\qquad\qquad \Rightarrow q_{HL}\,\theta_H - p_{HL} \geq q_{LL}\,\theta_L - p_{LL}
\end{aligned}
$$

Hence the binding IR_{LL} gives p_{LL}. To obtain p_{LH} substitute $q_{LL}\theta_L - p_{LL} = 0$ on the left hand side of (B.32).

We next want to further simplify the problem by eliminating redundant IC constrains. First, note that (B.31) ensures that (θ_H, θ_H) will not deviate to (θ_H, θ_L) and visa versa. Equation (B.32) implies the same for (θ_L, θ_H) and (θ_L, θ_L).

Next consider the incentives of the two period 1 high types (θ_H, θ_L) and (θ_H, θ_H) to deviate to either of the two period 1 low types (θ_L, θ_L) and (θ_L, θ_H). The derivations below show that if (θ_H, θ_H) prefers not to do any of those two potential deviations, then the same is true for (θ_H, θ_L).

$$\text{IC(HH,LL)} \;\; q_{HH}\theta_H + \mathbb{E}_H - p_{HH} \;\geq\; q_{LL}\theta_H + \mathbb{E}_L - p_{LL}$$
$$\overset{\rho_H \geq \rho_L}{\Longrightarrow} \;\; q_{HH}\theta_H - p_{HH} \;\geq\; q_{LL}\theta_H - p_{LL} \;\; \overset{\text{(B.31)}}{\Longrightarrow} \;\; q_{HL}\theta_H - p_{HL} \;\geq\; q_{LL}\theta_H - p_{LL}$$
$$\text{IC(HH,LH)} \;\; q_{HH}\theta_H + \mathbb{E}_H - p_{HH} \;\geq\; q_{LH}\theta_H + \mathbb{E}_L - p_{LH}$$
$$\overset{\rho_H \geq \rho_L}{\Longrightarrow} \;\; q_{HH}\theta_H - p_{HH} \;\geq\; q_{LH}\theta_H - p_{LH} \;\; \overset{\text{(B.31)}}{\Longrightarrow} \;\; q_{HL}\theta_H - p_{HL} \;\geq\; q_{LH}\theta_H - p_{LH}$$

Hence IC(HH,LL) and IC(HH,LH) imply IC(HL,LL) and IC(HL,LH), respectively. The first part of the first inequality given in the statement of the Lemma follows immediately from IC(HH,LL), while the first pat of the second inequality follows after substituting in IC(HH,LH) the price p_{LH} using (B.32).

Finally consider the incentives of the two period 1 low types (θ_L, θ_L) and (θ_L, θ_H) to deviate to either of the two period 1 low types (θ_H, θ_L) and (θ_H, θ_H). The derivations below show that if (θ_L, θ_L) prefers not to do any of those two potential deviations, then the same is true for (θ_L, θ_H).

$$\text{IC(LL,HL)} \;\; q_{LL}\theta_L - p_{LL} \;\geq\; q_{HL}\theta_L - p_{HL}$$
$$\overset{\text{(B.32)}}{\Longrightarrow} \;\; q_{LH}\theta_L - p_{LH} \;\geq\; q_{HL}\theta_L - p_{HL} \;\; \overset{\rho_H \geq \rho_L}{\Longrightarrow} \;\; q_{LH}\theta_L + \mathbb{E}_L - p_{LH} \;\geq\; q_{HL}\theta_L + \mathbb{E}_H - p_{HL}$$
$$\text{IC(LL,HH)} \;\; q_{LL}\theta_L - p_{LL} \;\geq\; q_{HH}\theta_L - p_{HH}$$
$$\overset{\text{(B.32)}}{\Longrightarrow} \;\; q_{LH}\theta_L - p_{LH} \;\geq\; q_{HH}\theta_L - p_{HH} \;\; \overset{\rho_H \geq \rho_L}{\Longrightarrow} \;\; q_{LH}\theta_L + \mathbb{E}_L - p_{LH} \;\geq\; q_{HH}\theta_L + \mathbb{E}_H - p_{HH}$$

Substituting zero, which is the payoff of (θ_L, θ_L) on the left hand side of IC(LL,HH) gives that $p_{HH} \geq q_{HH}\theta_H$, whereas the same substitution on IC(LL,HL) gives

$$p_{HL} \;\geq\; q_{HL}\theta_H \;\; \overset{\text{(B.31)}}{\Longrightarrow} \;\; p_{HH} \geq q_{HH}\theta_H$$

which completes the proof. □

But note that the only effect of sending an informative signal in ($\mathcal{P}_4$) is that it decreases the upper bound of p_{HH} and as a result of p_{HL}. Hence an informative signal is never strictly optimal. This result is due to the inclusion of (θ_L, θ_L), which makes it impossible for S_1 to charge (θ_L, θ_H) for the possibility of obtaining a discount.

We continue by considering the possibility of excluding (θ_L, θ_L) from S_1's contract.

Lemma B.14. *Suppose that S_1 uses a contract that excludes (θ_L, θ_L). Then S_2 charges price $\overline{p} = \theta_H$ irrespectively of the choice of $g(s \,|\, \theta_1)$.*

Proof. When (θ_L, θ_L) is excluded the lowest possible posterior on θ_2, with a signal that only depends on θ_1, is achieved when S_1 reveals the buyer as period 1 high type. In this case

$$\mu_2^s \;=\; \rho_H \;\geq\; \frac{\theta_L}{\theta_H}$$

from which the statement of the lemma follows. □

Therefore, again in this case no information provision is optimal, and since excluding (θ_L, θ_L) achieves nothing S_1 is no better off compared to the four type contract.

Another possibility that is potentially incentive compatible is to exclude both (θ_L, θ_L) and (θ_H, θ_L), but in this case it is trivial to argue that again S_2 will charge price $\overline{p} = \theta_H$ irrespectively of the choice of $g(s \,|\, \theta_1)$.

Hence we have considered all cases and we have shown that no information provision will always be optimal.

C Isoelastic Cost

In this section we expand the baseline model by allowing q_t to take any positive value, but we introduce an isoelastic cost function. Hence the payoff of each seller is

$$\pi_t \;=\; p_t - c(q_t)\,, \quad \text{where} \quad c(q) \;=\; \frac{q^{1+\frac{1}{\epsilon}}}{1+\frac{1}{\epsilon}} \quad \text{for} \;\; \epsilon \;>\; 0\,,$$

and we interpret p_t as the total price charged by S_t for all of the traded quantity q_t. The buyer's payoff from each trade is

$$\theta_1\, q_1 - p_1 \qquad \text{and} \qquad b\, \theta_2\, q_2 - p_2\,,$$

when served by S_1 and S_2, respectively. The constant $b > 0$ is introduced to allow for the possibility that the buyer's valuation of each trade can vary in some deterministic way. To maintain the analysis as close to the baseline model as possible we assume that

$$\max\left\{\rho_L\,,\,\frac{\theta_L}{\theta_H}\right\}\;\leq\;\rho_H \tag{C.1}$$

which ensures (i) that the agent's type is positively correlated across sellers, and (ii) that if S_1 reveals the buyer as a high period 1 type, then S_2 will only supply a positive quantity to the period 2 high type. Finally, to demonstrate that controlling the flow of information can be payoff equivalent for S_1 to controlling access to S_2, we assume instead that S_1 simple gets to interact always first with the buyer and can commit on a single distribution even in the absence of a contract. Hence even if the buyer opts not to trade with her.

The analysis follows closely that of the baseline model. We first solve S_2's payoff maximisation problem and derive the buyer's information rents from his second contract. Subsequently, this is used to build and solve S_1's information provision problem in the first and second best, in subsections C.2 and C.3, respectively.

C.1 The buyer's post contractual payoff and outside option

We start by solving S_2's payoff maximisation problem for any given realisation of her posterior on the period 1 type μ_1^s, and period 2 type μ_2^s. As before, those two satisfy

$$\mu_2^s\;=\;\mu_1^s\,\rho_H+(1-\mu_1^s)\,\rho_L$$

S_2's problem is similar to that of the baseline model and its treatment can be found in the appendix. In the following lemma we provide the buyer's payoff, which is the only result needed to proceed with S_1's problem.

Lemma C.1. *The payoff of a low buyer type under S_2 is equal to zero, while that of the high one equals*

$$Q(\mu_1^s)\;=\;\begin{cases} b^{1+\epsilon}\cdot(\theta_H-\theta_L)\cdot\left(\frac{\theta_L-\mu_2^s\theta_H}{1-\mu_2^s}\right)^{\epsilon} & ,\ \mathit{if}\ \ \mu_1^s\leq\mu^* \\ 0 & ,\ \mathit{if}\ \ \mu_1^s\geq\mu^* \end{cases} \tag{C.2}$$

Also, on the subset of posteriors $[0,\mu^*)$ it is decreasing, and strictly concave (convex) for

$$\mu_1^s\quad<(>)\quad\mu^*\,+\,\frac{1-\epsilon}{2\cdot(\rho_H-\rho_L)}\cdot\left(1-\frac{\theta_L}{\theta_H}\right)$$

Proof. Follows from the corresponding lemma of the section of multi-period contracts on the online appendix once you substitute $Q(\mu_1^s)$ with $B(\mu_2^s)$. □

Therefore, only a high period 2 type obtains a positive payoff under S_2, as in the baseline model. But contrary to it, his payoff is a continuous function of S_2's posterior. Interestingly, for $\epsilon \to 0$ the isoelastic model converges to the baseline one.

Next, we want to demonstrate how S_1 can enforce the buyer's participation in her contract by using the event $\{\emptyset\}$. This denotes a rejection of S_1's offer from the buyer, and we will show that it will not occur on the equilibrium path. We have assumed that S_1 can commit on the signal distribution g even if the buyer does not participate in her contract[6] . Hence, let $g(s \mid \emptyset)$ be the corresponding conditional distribution. Therefore, the outside options of a period 1 high and low type are

$$\rho_H \mathbb{E}_g \left[Q(\mu_1^s) \mid \emptyset \right] \quad \text{and} \quad \rho_L \mathbb{E}_g \left[Q(\mu_1^s) \mid \emptyset \right]$$

respectively. Thus, S_1's objective is to reduce $\mathbb{E}_g \left[Q(\mu_1^s) \mid \emptyset \right]$ as much as possible. This can be made zero by introducing signal realisation $s_\emptyset \in S$ such that

$$\left. \begin{array}{r} g(s_\emptyset \mid \emptyset\ ) = 1 \\ g(s_\emptyset \mid \theta_L\,) = 0 \\ g(s_\emptyset \mid \theta_H) = \varepsilon \end{array} \right\} \quad \Rightarrow \quad \Pr(\theta_1 = \theta_H \mid s_\emptyset) \;=\; \frac{\mu_0\, \varepsilon}{\mu_0\, \varepsilon + \Pr(\emptyset)}$$

Suppose S_1 was using a mechanism such that $\Pr(\emptyset)$=0, that is the buyer was always participating on path. Then $\Pr(\theta_1 = \theta_H \mid s_\emptyset) = 1$ and this would be true even if ε was an infinitesimal positive real number. Effectively, the buyer's outside option under such a mechanism would be zero as

$$\mathbb{E}_g \left[Q(\mu_1^s) \mid \emptyset \right] \;=\; Q(1) \;=\; 0$$

In addition, by setting $\varepsilon \to 0$ the cost of including $s_\emptyset$ on the set of possible realisations S tends to zero, because the same is true for the probability of using it on path.

Hence there exists a mechanism that achieves the smallest possible outside option for the buyer, and the cost of doing so is zero since the effect of $s_\emptyset$ on S_1's information provision problem is infinitesimal. Moreover, the signal $s_\emptyset$ can be used in both the first and second best. Interestingly, this make the model equivalent to one in which S_1 controls not only the information provided to S_2, but also the buyer's access to her. Hence to save in space, and because $s_\emptyset$ does not occur on path, we will hereafter ignore it on both the proofs and statements of our results, and we will set the buyer's outside option immediately to zero.

[6]The discussion here demonstrates how S_1 can reduce the buyer's outside option to zero. In the baseline model, even if S_1 was not able to commit without a contract on g, the solution of her information provision problem would be the same. To see this note that the argumentation on Section 2 does not rely on the architect controlling access to the designer.

C.2 S_1's first best contract

We solve S_1's revenue maximisation problem under the assumption that if the buyer opts to participate in her mechanism, then his type is automatically reveal to her, but not to S_2. Hence she solves

$$\begin{aligned} \max_{p_1,q_1,g} \quad & \mu_0\Big(p_1(\theta_H) - c\big[q_1(\theta_H)\big]\Big) + (1-\mu_0)\Big(p_1(\theta_L) - c\big[q_1(\theta_L)\big]\Big) \\ \text{s.t. } (\mathrm{IR}_L) \quad & \theta_L\, q_1(\theta_L) - p_1(\theta_L) + \rho_L\, \mathbb{E}_g\left[Q(\mu_1^s)\,|\,\theta_L\right] \;\geq\; 0 \\ (\mathrm{IR}_H) \quad & \theta_H q_1(\theta_H) - p_1(\theta_H) + \rho_H \mathbb{E}_g\left[Q(\mu_1^s)\,|\,\theta_H\right] \;\geq\; 0 \end{aligned} \tag{$\mathcal{P}_f$}$$

Both of the individual rationality constraints need to bind. Hence solve for the prices, and substitute those in S_1's objective function to obtain the unconstrained problem

$$\max_{q_1,g} \quad \left\{ \begin{array}{c} \mu_0\cdot\Big(\theta_H q_1(\theta_H) - c\big[q_1(\theta_H)\big]\Big) + (1-\mu_0)\cdot\Big(\theta_L q_1(\theta_L) - c\big[q_1(\theta_L)\big]\Big) \\ +\,\mu_0\rho_H\, \mathbb{E}_g\left[Q(\mu_1^s)\,|\,\theta_H\right] + (1-\mu_0)\rho_L \mathbb{E}_g\left[Q(\mu_1^s)\,|\,\theta_L\right] \end{array} \right\} \tag{$\mathcal{P}_f'$}$$

The first line represents the surplus generated from the trade of period 1, while the second the buyer's ex ante post contractual payoff, which S_1 captures through the individual rationality constrains. We can use first order conditions to obtain that the optimal supply schedule in the first best is

$$q_f(\theta_1) \;=\; (\theta_1)^{\epsilon}$$

In the remaining of this subsection we focus on S_1's information provision problem

$$\max_{g} \quad \Big\{\mu_0\rho_H\, \mathbb{E}_g\left[Q(\mu_1^s)\,|\,\theta_H\right] + (1-\mu_0)\rho_L \mathbb{E}_g\left[Q(\mu_1^s)\,|\,\theta_L\right]\Big\} \tag{$\mathcal{G}_f$}$$

Lemma C.2. *In the first best, S_1's information provision problem equivalently becomes*

$$\max_{g} \;\; \mathbb{E}_g[J_f(\mu_1^s)] \tag{$\mathcal{G}_f$}$$

where its point-wise value $J_f(\mu_1^s)$ is

$$J_f(\mu_1^s) \;=\; Q(\mu_1^s)\cdot\left[\,\mu_1^s\cdot(\rho_H - \rho_L) + \rho_L\,\right] \tag{C.3}$$

Proof. Identical to that of Lemma 2. ☐

Then following the same argumentation as in the main text we get that S_1's information provision problem becomes

$$\max_{\tilde{g}} \ \mathbb{E}_{\tilde{g}}\big[J_f(\mu)\big] \quad \text{s.t.} \quad \mathbb{E}_{\tilde{g}}[\,\mu\,] \;=\; \mu_0 \tag{$\mathcal{G}'_f$}$$

To solve this it is important to characterise the graph of $J_f(\mu)$. Given our existing assumption (C.1), the only case in which information provision can have an impact on the mechanism used by S_2 is when

$$\rho_L \;<\; \frac{\theta_L}{\theta_H} \;<\; \rho_H \tag{C.4}$$

in which case $\mu^* \in (0,1)$. Also, define

$$\mu_f^{**} \;=\; \max\left\{0\,,\, \frac{\beta_f^{**}-\rho_L}{\rho_H-\rho_L}\right\} \quad \text{where} \quad \beta_f^{**} \;=\; \frac{2\,\theta_L/\theta_H}{2+(\epsilon-1)\big(1-\frac{\theta_L}{\theta_H}\big)}$$

Lemma C.3. *Suppose that (C.4) holds. Then $J_f(\mu)$ is positive on $[0,\mu^*)$ and equals zero on $[\mu^*,1]$. Moreover,*

- *It changes monotonicity at most once*
- *If $\epsilon \leq 1$, then it is strictly concave on $[0,\mu^*]$.*
- *If $\epsilon > 1$, then it is strictly concave in $[0,\mu_f^{**}]$, and strictly convex in $[\mu_f^{**},\mu^*]$.*

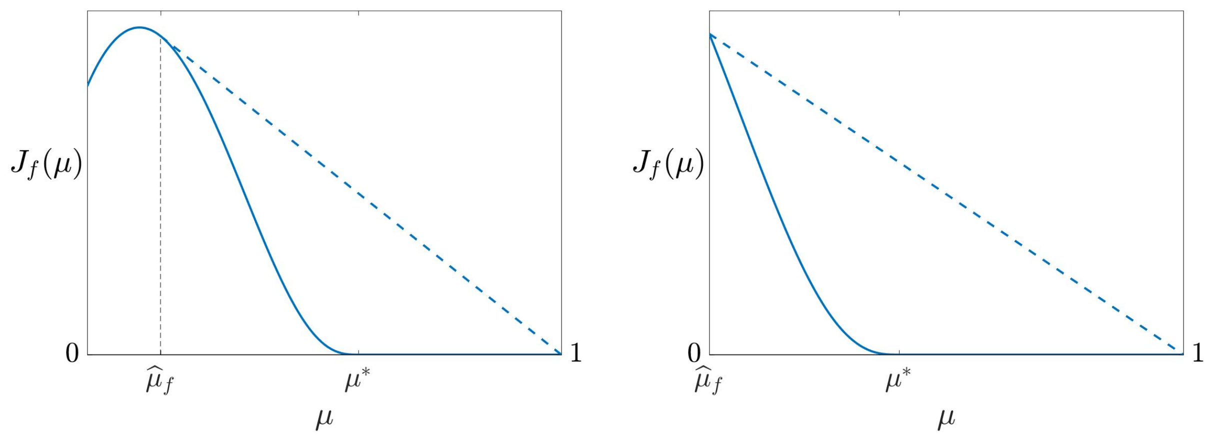
Proof. Follows as a subcase of Lemma 5.2 of the online appendix. In particular, the functional form of $J_f(\mu)$ is the same with that of $J_t(\mu,\lambda)$ on the $(\lambda = 0)$ boundary. □

Plots (8a) and (8b) demonstrate the two possible cases of J_f under (C.4). In the former plot it is initially increasing and subsequently decreasing, whereas in the latter it is only decreasing. An intuitive way to understand the shape of J_f is to consider its value on the prior μ_0, where the underline trade off between $Q(\mu_0)$ and $\mu_0\rho_H + (1-\mu_0)\rho_L$ becomes apparent. That is between the rents captured by a period 2 high type and the probability of the buyer to be one. The higher the μ_0 is, the smaller the $Q(\mu_0)$, which has a negative impact on S_1's post contractual payoff. However, only a period 2 high type captures positive rents, which creates an effect opposite from the above.

As in the baseline model, we solve S_1's information provision problem ($\mathcal{G}'_f$) by invoking the optimality condition $\mathbb{E}_{\tilde{g}}[J_f(\mu)] = \mathcal{J}_f(\mu_0)$, where

$$\mathcal{J}_f(\mu) \;=\; \sup\big\{z\,|\,(\mu,z)\in \text{co}(J_f)\big\},$$

denotes the concave closure of J_f. In Figure (8) whenever $\mathcal{J}_f(\mu) > J_f(\mu)$ this is represented by the dashed line. If $\mu < \hat{\mu}_f$, then there is not a linear combination of points of J_f that



(a) small ρ_L (b) big ρ_L

Figure 8: A representative graph of J_f. The dashed line denotes its concave closure, when this is above J_f.

achieves something above $J_f(\mu)$. Henceforth, on this set of points $J_f(\mu) = \mathcal{J}_f(\mu)$ and there is no benefit from information provision. On the other hand, for $\mu > \hat{\mu}_f$ finding such a linear combination is possible. In fact, in this case $\mathcal{J}_f(\mu)$ is the line that connects $J_f(1)$ to the tangency point $\hat{\mu}_f$, which is the unique solution of

$$J_f(\hat{\mu}_f) + J_f'(\hat{\mu}_f)(1 - \hat{\mu}_f) \;=\; 0 \tag{C.5}$$

Interestingly, the tangency point is not necessarily in $[0, 1]$, as demonstrated in Plot (8b), in which case we instead use the corner solution $\hat{\mu}_f = 0$.

Proposition C.1. *In the first best, an informative signal strictly solves S_1's payoff maximisation problem ($\mathcal{P}_f$) iff (C.4) holds and $\mu_0 > \hat{\mu}_f$. If those two conditions hold, then an optimal signal is $s \in \{\underline{s}, \overline{s}\}$ with distribution*

$$g_f\left(\underline{s}\,|\,\theta_L\right) \;=\; 1\,, \qquad \textit{and} \qquad g_f\left(\underline{s}\,|\,\theta_H\right) \;=\; \frac{1-\mu_0}{\mu_0}\frac{\hat{\mu}_f}{1-\hat{\mu}_f}. \tag{C.6}$$

In addition, the optimal supply schedule is $q_f(\theta_1) \;=\; (\theta_1)^{\epsilon}$.

Proof. The above discussion implies that the concave closure of $J_f(\mu)$ is given by

$$\mathcal{J}_f(\mu) \;=\; \begin{cases} J_f(\mu) & \text{, for } \;\mu \;\leq\; \hat{\mu}_f \\ J_f(\hat{\mu}_f) + J_f'(\hat{\mu}_f)(\mu - \hat{\mu}_f) & \text{, for } \;\mu \;\geq\; \hat{\mu}_f \end{cases} \tag{C.7}$$

The functional form of $\hat{\mu}_f$ follows as a subcase of Proposition 1.2 when $\Psi_i = 0$. This is

$$\hat{\mu}_f \;=\; \max\left\{0\,,\,\frac{\hat{\beta}_f-\rho_L}{\rho_H-\rho_L}\right\} \quad \text{where} \quad \hat{\beta}_i \;=\; \frac{\omega_1-\sqrt{(\omega_1)^2-4\omega_0\omega_2}}{2\omega_2}$$

and

$$\omega_0 \;=\; \frac{\theta_L}{\theta_H}\,, \quad \omega_1 \;=\; 1+\frac{\theta_L}{\theta_H}+\epsilon\cdot\left(1-\frac{\theta_L}{\theta_H}\right) \quad \omega_2 \;=\; 1+\frac{\epsilon}{\rho_H}\cdot\left(1-\frac{\theta_L}{\theta_H}\right) \tag{C.8}$$

If $\mu_0 \le \hat{\mu}_f$, then $J_f(\mu_0) = \mathcal{J}_f(\mu_0)$ and as a result setting $\Pr(\mu = \mu_0) = 1$ is optimal, which is achieved when no information is provided to S_2. If $\mu_0 > \hat{\mu}_f$, then the optimal signal randomises between posteriors $\hat{\mu}_f$ and 1. This is the solution of ($\mathcal{G}'_f$). Hence if $s \in \{\underline{s}, \overline{s}\}$ solves ($\mathcal{G}_f$), then it has to be that $g_f\,(\overline{s}\,|\,\theta_L) = 0$, so that the posterior that $\overline{s}$ implies is one. Therefore, $g_f\,(\underline{s}\,|\,\theta_L) = 1$. Finally, to find $g_f\,(\underline{s}\,|\,\theta_H)$ note that this has to satisfy

$$\hat{\mu}_f \;=\; \frac{\mu_0\, g_f\,(\underline{s}\,|\,\theta_H)}{\mu_0\, g_f\,(\underline{s}\,|\,\theta_H)+1-\mu_0} \tag{C.9}$$

But if this signal solves ($\mathcal{G}_f$), then it also solves ($\mathcal{P}'_f$), and as a result ($\mathcal{P}_f$). □

C.3 S_1's second best contract

Next we analyse the second best, where θ_1 is the buyer's private information. S_1 solves

$$\begin{aligned}
&\max_{p_1(\widehat{\theta}_1),\, q_1(\widehat{\theta}_1),\, g(s\,|\,\widehat{\theta}_1)} \quad \Big\{\, \mu_0\Big(p_1(\widehat{\theta}_H)-c\big[q_1(\widehat{\theta}_H)\big]\Big)+(1-\mu_0)\,\Big(p_1(\widehat{\theta}_L)-c\big[q_1(\widehat{\theta}_L)\big]\Big)\,\Big\} \\
&\text{s.t. } (\text{IR}_L),\;\; (\text{IR}_H)\,, \\
&(\text{IC}_L)\;\; \theta_L\, q_1(\widehat{\theta}_L)\,+\rho_L\,\mathbb{E}_g\left[Q(\mu_1^s)\,|\,\widehat{\theta}_L\right]-p_1(\widehat{\theta}_L) \\
&\qquad\qquad \ge\; \theta_L\, q_1(\widehat{\theta}_H)\,+\rho_L\,\mathbb{E}_g\left[Q(\mu_1^s)\,|\,\widehat{\theta}_H\right]-p_1(\widehat{\theta}_H) \\
&(\text{IC}_H)\;\; \theta_H q_1(\widehat{\theta}_H)+\rho_H\mathbb{E}_g\left[Q(\mu_1^s)\,|\,\widehat{\theta}_H\right]-p_1(\widehat{\theta}_H) \\
&\qquad\qquad \ge\; \theta_H\, q_1(\widehat{\theta}_L)\;+\rho_H\mathbb{E}_g\left[Q(\mu_1^s)\,|\,\widehat{\theta}_L\right]-p_1(\widehat{\theta}_L)
\end{aligned} \tag{$\mathcal{P}$}$$

where the individual rationality constraints, (IR_L) and (IR_H), are as in the previous subsection. Hereafter, we will use $\{p_L, p_H, q_L, q_H\}$ instead of $\{p_1(\widehat{\theta}_L), p_1(\widehat{\theta}_H), q_1(\widehat{\theta}_L), q_1(\widehat{\theta}_H)\}$, in order to maintain a compact notation. Similarly to the first best, it is convenient to reduce the number of constrains by substituting the transfers p_L and p_H.

Lemma C.4. *In the second best, S_1's payoff maximisation problem equivalently becomes*

$$\max_{q_1,g} \quad \Bigg\{ \mu_0 \cdot \Big(\theta_H q_H - c(q_H)\Big) + (1-\mu_0) \cdot \left(\frac{\theta_L - \mu_0 \theta_H}{1-\mu_0} \cdot q_L - c(q_L)\right) + \mu_0 \rho_H \, \mathbb{E}_g \left[Q(\mu_1^s) \,|\, \theta_H\right] + (1-\mu_0)\rho_L \mathbb{E}_g \left[Q(\mu_1^s) \,|\, \theta_L\right] - \mu_0(\rho_H - \rho_L)\mathbb{E}_g \left[Q(\mu_1^s) \,|\, \theta_L\right] \Bigg\} \quad (\mathcal{P}')$$

$$s.t \quad (\theta_H - \theta_L)\,(q_H - q_L) \;\geq\; (\rho_H - \rho_L) \Big(\mathbb{E}_g\big[Q(\mu_1^s) \,|\, \theta_L\big] - \mathbb{E}_g\big[Q(\mu_1^s) \,|\, \theta_H\big] \Big) \qquad (\mathcal{P}_c)$$

Proof. Identical to that of Lemma 3. □

The first line represents the surplus of the first period trade that S_1 is able to capture. Similarly, the second line is the part of the buyer's ex ante payoff that S_1 captures. Both lines are smaller than the corresponding ones of ($\mathcal{P}'_f$), because of the period 1 high type's rents. The point-wise optimal production level is

$$q_1^*(\theta_1) \;=\; \begin{cases} (\theta_H)^\epsilon & , \text{if } \theta_1 = \theta_H \\ (\xi)^\epsilon & , \text{if } \theta_1 = \theta_L \end{cases} , \quad \text{where } \xi \;=\; \max\left\{ 0 \,,\, \theta_L - (\theta_H - \theta_L) \cdot \frac{\mu_0}{1-\mu_0} \right\} \quad (C.10)$$

However, because of ($\mathcal{P}_c$) we will not always be able to find a contract that implements this supply schedule. We will say that a signal is *impactful* if $\mathbb{E}_g\big[Q(\beta_0^s) \,|\, \theta_L\big] > \mathbb{E}_g\big[Q(\beta_0^s) \,|\, \theta_H\big]$. As we have argued in Remark 1 whenever a signal influences S_2's contract, this inequality has to hold.

Lemma C.5 (Implementation)**.** *For $\rho_H > \rho_L$:*

- *For any $q_1(\theta_1)$ combined with an impactful signal, there exists a constant b high enough for those to not be implementable.*
- *Conversely, the point-wise optimal $q_1^*(\theta_1)$ combined with an uninformative signal is implementable.*
- *A sufficient condition for $q_1^*(\theta_1)$ combined with any signal to be implementable is that*

$$\left(\frac{\theta_H}{\theta_L}\right)^\epsilon \;\geq\; 1 + (\rho_H - \rho_L) \cdot b^{1+\epsilon} \qquad (C.11)$$

Proof. To obtain the first statement note that ($\mathcal{P}_c$) can equivalently be re-written as

$$q_1(\theta_H) - q_1(\theta_L) \;\geq\; (\rho_H - \rho_L)\, b^{1+\epsilon} \times \left(\mathbb{E}_g \left[\left(\frac{\theta_L - \theta_H \mu_2^s}{1-\mu_2^s}\right)^\epsilon \,\middle|\, \theta_L \right] - \mathbb{E}_g \left[\left(\frac{\theta_L - \theta_H \mu_2^s}{1-\mu_2^s}\right)^\epsilon \,\middle|\, \theta_H \right] \right)$$

Hence as long as the right hand side is positive, it can be made infinitely large by increasing b. The second statement follows trivially from the above expression, as in the case of an uninformative signal the two expectations on the second line of its right hand side are equal to each other. Thus this becomes zero. For the third statement note that the left hand side of the above expression is bigger than $(\theta_H)^\epsilon - (\theta_L)^\epsilon$. Also, the second line of its right hand side is smaller than $\mathbb{E}_s\big[Q(\beta_0^s)\,|\,\theta_L\big] \leq Q(0) = (\theta_L)^\epsilon$. Therefore ($\mathcal{P}_c$) is implied by

$$(\theta_H)^\epsilon - (\theta_L)^\epsilon \;\; \geq \;\; b^{1+\epsilon}\,(\rho_H - \rho_L)\,(\theta_L)^\epsilon\,,$$

which after rearranging gives (C.11). □

The lemma sheds light on the limitation of point-wise maximisation, for both production and information provision, in this setup. To be more precise, when the buyer's report has post contractual value, in the sense that an informative signal is optimal, S_1 may not always be able to provide the right incentives for the buyer to share this information. In particular, the first statement above shows that when the interaction of period 2 is much more important than that of period 1, that is b is relatively big or equivalently θ_L and θ_H are relatively small, S_1 may have to significantly constrain the amount of information she shares with S_2. In fact if $b \to \infty$, then on this limiting case S_1 cannot provide any information at all. Dworczak (2016*a*,*b*) has demonstrated the same limitation by showing in a setting closer to an action that the only type of mechanism that is always implementable, under an aftermarket for the bidder that acquires the object, is one that does not share information on his reported type. On the other hand, when no information provision is optimal, as in Calzolari and Pavan (2006), this is always implementable.

The rest of the analysis will focus on deriving the point-wise optimal signal. First, because for the right choice of b, or θ_H/θ_L, condition (C.11) will always be satisfied. Second, because for this case the graphical approach used in the previous subsection still applies, which makes the solution of S_1's information provision problem much more tractable and ease to demonstrate[7]. S_1's information provision problem, ignoring the implementation constrain (C.11), becomes

$$\max_g \;\Big\{ \mu_0\rho_H\,\mathbb{E}_g\left[Q(\mu_1^s)\,|\,\theta_H\right] + (1-\mu_0)\rho_L\mathbb{E}_g\left[Q(\mu_1^s)\,|\,\theta_L\right] \,-\, \mu_0(\rho_H-\rho_L)\mathbb{E}_g\left[Q(\mu_1^s)\,|\,\theta_L\right] \Big\} \quad (\mathcal{G})$$

[7]Despite that, the implications of a binding implementability constrain pose a very interesting question, worthy of further research.

Lemma C.6. *In the second best, S_1's information provision problem equivalently becomes*

$$\max_{g} \ \mathbb{E}_g[J(\mu_1^s)] \tag{$\mathcal{G}$}$$

where its point-wise value $J(\mu_1^s)$ is

$$J(\mu_1^s) \ = \ \frac{\rho_H - \rho_L}{1-\mu_0} \cdot Q(\mu_1^s) \cdot \left(\mu_1^s - \frac{\mu_0\rho_H - \rho_L}{\rho_H - \rho_L}\right) \tag{C.12}$$

Proof. Identical to that of Lemma 4. □

Similarly to the baseline model ($\mathcal{G}$) is transformed to a choice of distributions over posteriors $\tilde{g}(\mu)$, that is it equivalently becomes

$$\max_{\tilde{g}} \ \mathbb{E}_{\tilde{g}}\big[J(\mu)\big] \quad \text{s.t.} \ \ \mathbb{E}_{\tilde{g}}[\,\mu\,] \ = \ \mu_0 \tag{$\mathcal{G}'$}$$

This is solved by invoking the optimality condition $\mathbb{E}_{\tilde{g}}\big[J(\mu)\big] = \mathcal{J}(\mu_0)$, where

$$\mathcal{J}(\mu) \ = \ \sup\big\{z \,|\, (\mu, z) \in \text{co}(J_0)\big\}$$

denotes the concave closure of J. To find $\mathcal{J}$ the shape of J needs to be characterised, which is undertaken in the following Lemma. To shorten its statement define

$$\mu^{**} = \max\left\{0 \ , \ \frac{\beta^{**} - \rho_L}{\rho_H - \rho_L}\right\} \ , \quad \text{and} \quad \beta^{**} \ \equiv \ \frac{\theta_L}{\theta_H} \, \frac{2\big(1-\mu_0\rho_H\big) + (\epsilon - 1)\big(1 - \frac{\theta_L}{\theta_H}\big)\frac{\mu_0\rho_H}{\theta_L/\theta_H}}{2(1-\mu_0\rho_H) + (\epsilon-1)\big(1-\frac{\theta_L}{\theta_H}\big)} \, .$$

As in the first best, we only need to consider the shape of J when (C.4) holds, since otherwise information provision has no impact on S_2's contract.

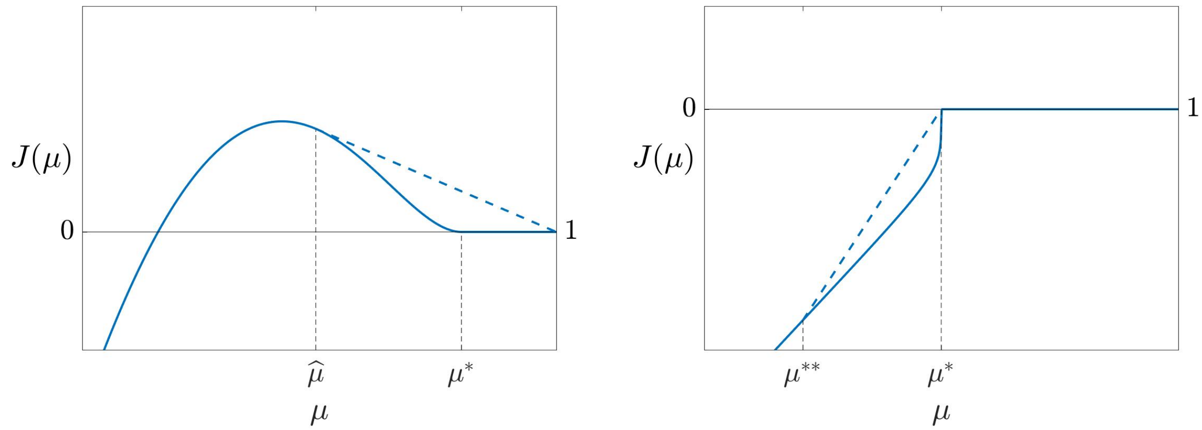
Lemma C.7. *Assume throughout that (C.4) holds. Suppose $\mu_0\rho_H < \frac{\theta_L}{\theta_H}$, then $J(\mu)$:*

- *changes monotonicity at most once, and if $\mu^* > 0$ it falls to it from above.*
- *If $\epsilon \leq 1$, then it is strictly concave on $[0, \mu^*]$.*
- *If $\epsilon > 1$, then it is strictly concave in $[0, \mu^{**}]$, and strictly convex in $[\mu^{**}, \mu^*]$.*

If instead $\mu_0\rho_H \geq \frac{\theta_L}{\theta_H}$, then $J(\mu)$ is non-positive.

Proof. Follows as a subcase of Lemma 5.2 in the online appendix. □

Figure (9) shows the two possible cases of $J(\mu)$. That first one, shown in Plot (9a), is qualitatively similar to the graph of J_f from the previous subsection. The second, presented



(a) $\mu_0\rho_H < \frac{\theta_L}{\theta_H}$ (b) $\mu_0\rho_H \geq \frac{\theta_L}{\theta_H}$

Figure 9: A representative graph of J. The dashed line denotes its concave closure, when this is above J.

in Plot (9b), is substantially different, and in particular in non-positive. This is because if $\mu_0\rho_H$ is relatively high, then convincing the period 1 high type to report truthfully, under an informative signal, becomes too expensive. Similarly, to the baseline model.

Hence if $\rho_H\mu_0 < \theta_L/\theta_H$, then the analysis is identical to that of the first best, that is the dashed line on Plot (9a) represents $\mathcal{J}(\mu)$ for $\mu > \hat{\mu}$, and $\hat{\mu}$ is found by solving the tangency condition

$$J(\hat{\mu}) + J'(\hat{\mu})(1-\hat{\mu}) \;=\; 0 \tag{C.13}$$

On the other hand, if $\rho_H\mu_0 \geq \theta_L/\theta_H$, then no information provision has to be optimal since it always achieves at least zero.

Proposition C.2. *Suppose throughout that the implementation condition* (C.11) *holds. Then the point-wise optimal supply schedule, as given in* (C.10)*, solves S_1's payoff maximisation problem* ($\mathcal{P}$)*. In addition, an informative signal strictly solves* ($\mathcal{P}$) *iff*

$$\max\{\rho_L,\, \rho_H\mu_0\} \;<\; \frac{\theta_L}{\theta_H} \;<\; \rho_H\,, \qquad \textit{and} \qquad \mu_0 > \hat{\mu} \tag{C.14}$$

If those two conditions hold, then an optimal signal is $s \in \{\underline{s}, \overline{s}\}$ with distribution

$$g^*\left(\underline{s}\,|\,\theta_L\right) \;=\; 1\,, \qquad \textit{and} \qquad g^*\left(\underline{s}\,|\,\theta_H\right) \;=\; \frac{1-\mu_0}{\mu_0}\frac{\hat{\mu}}{1-\hat{\mu}} \tag{C.15}$$

Proof. Whenever (C.11) holds, the point-wise optimal supply schedule is implementable under any signal. Hence the solution of ($\mathcal{G}$) will also solve ($\mathcal{P}'$), and as a result ($\mathcal{P}$).

Then we can focus on finding the $\tilde{g}$ that solves ($\mathcal{G}'$). It follows from the above discussion that whenever $\mu_0\rho_H \geq \theta_L/\theta_H$ no information provision is optimal. If instead $\mu_0\rho_H < \theta_L/\theta_H$, but (C.4) does not hold, then information provision has no impact on S_2's contract, in which case no information provision is again optimal.

Hence it remains to consider $\mu_0\rho_H < \theta_L/\theta_H$ when (C.4) holds. As we argued before in this case the concave closure of J is

$$\mathcal{J}(\mu) = \begin{cases} J(\mu) & , \text{for } \mu \leq \hat{\mu} \\ J(\hat{\mu}) + J'(\hat{\mu})(\mu - \hat{\mu}) & , \text{for } \mu \geq \hat{\mu} \end{cases} \tag{C.16}$$

where the functional form of the tangency point $\hat{\mu}$ follows as a subcase of Proposition 1.2 in the online appendix when $\Psi_i = \mu_0\rho_H$.

$$\hat{\mu} = \max\left\{0\,, \frac{\hat{\beta} - \rho_L}{\rho_H - \rho_L}\right\} \quad \textit{where} \quad \hat{\beta} = \frac{\omega_1 - \sqrt{(\omega_1)^2 - 4\omega_0\omega_2}}{2\omega_2} \tag{C.17}$$

and

$$\omega_0 = \frac{\theta_L}{\theta_H} + \frac{\mu_0\rho_H k}{1-\mu_0}\,, \quad \omega_1 = 1 + \frac{\theta_L}{\theta_H} + k\frac{1+\mu_0}{1-\mu_0}\,, \quad \omega_2 = 1 + \frac{k}{\rho_H(1-\mu_0)}\,, \tag{C.18}$$

and $k = \epsilon\left(1 - \frac{\theta_L}{\theta_H}\right)$ is introduced to maintain a compact notation. Note that if the tangency point is negative, then we use the corner solution instead, which is zero. Hence we can conclude that information provision is strictly optimal iff $\mu_0 > \hat{\mu}$, in which case the solution of ($\mathcal{G}'$) randomises between $\hat{\mu}$ and one. Let $\overline{s}$ be the signal that results in $\mu = 1$. Then it has to be that $g^*(\overline{s}\,|\,\theta_L) = 0$. Therefore, the binary signal $s \in \{\underline{s}, \overline{s}\}$ solves ($\mathcal{G}$) if $g^*(\underline{s}\,|\,\theta_L) = 1$ and $g^*(\underline{s}\,|\,\theta_H)$ satisfies

$$\hat{\mu} = \frac{\mu_0\, g^*(\underline{s}\,|\,\theta_H)}{\mu_0\, g^*(\underline{s}\,|\,\theta_H) + 1 - \mu_0} \tag{C.19}$$

□

Compared to the baseline model, it is harder to infer the shape of the subset of transitioning probabilities for which information provision is strictly optimal. This is because instead of the linear constrain (12) that used to define the diagonal of the triangle that was this set for the baseline model, we now have the non-linear constrain $\mu \geq \hat{\mu}$. Despite that, we can still make the following claim. Define

$$(\hat{\rho}_L, \hat{\rho}_H) = \begin{cases} (\tilde{\rho}_L, 1) & , \text{if } \mu_0 < \theta_L/\theta_H \\ \left(0, \frac{\theta_L/\theta_H}{\mu_0}\right) & , \text{if } \mu_0 \geq \theta_L/\theta_H \end{cases}$$

where $\tilde{\rho}_L$ is the ρ_L that solves $\mu_0 + (1-\mu_0)\rho_L = \hat{\beta}$ when $\rho_H = 1$ and $\mu_0 < \theta_L/\theta_H$.

Corollary 3. *The set of points for which (C.14) is satisfied is a convex subset of*

$$\left\{(\rho_L, \rho_H) \;:\; \rho_L \;\in\; \left[\hat{\rho}_L\,,\,\frac{\theta_L}{\theta_H}\right] \;\; \textit{and} \;\; \rho_H \;\in\; \left[\frac{\theta_L}{\theta_H}\,,\,\hat{\rho}_H\right]\right\} \tag{C.20}$$

In addition, if for given point (ρ_L, ρ_H) information provision is strictly optimal in the second best, then the same is true for the first best.

For a graphical illustration of the above result also check Plot (6b). The comparison between the first and second best is of special interest. As shown the set of transitioning probabilities ρ_L and ρ_H for which information provision is optimal is larger in the former setup. This is due to the information rents captured by the period 1 high type, which create an additional cost that S_1 has to incur if she opts for an informative signal.

Proof . The functional form of $\hat{\mu}$ has being given in (C.17) and (C.18). It follows from the proof of Proposition 1.2 in the online appendix that $\hat{\beta}$ is the smaller of the two solutions of

$$\omega_2\beta^2 - \omega_1\beta + \omega_0 \;=\; 0 \tag{C.21}$$

with respect to β. Hence the left hand side of the above is zero when calculated on $\hat{\beta}$. If we calculate it on $\beta = \rho_H$ we get

$$\omega_2\rho_H^2 - \omega_1\rho_H + \omega_0 \;=\; \rho_H^2 - \rho_H\left(\frac{\theta_L}{\theta_H} + 1\right)\frac{\theta_L}{\theta_H} \;=\; (1-\rho_H)\left(\frac{\theta_L}{\theta_H} - \rho_H\right) \;\leq\; 0$$

Therefore, the left hand side of (C.21) is non-positive for $\beta \in [\hat{\beta}, \rho_H]$.

$$\mu_0 \;\geq\; \frac{\hat{\beta} - \rho_L}{\rho_H - \rho_L} \quad\Leftrightarrow\quad \underbrace{\mu_0\rho_H + (1-\mu_0)\rho_L}_{=\beta_0} \;\geq\; \hat{\beta}$$

But the biggest possible value of β_0 is ρ_H, hence we infer that

$$\mu_0 \;\geq\; \frac{\hat{\beta} - \rho_L}{\rho_H - \rho_L} \;\Leftrightarrow\; Q(\rho_H, \rho_L) \;\leq\; 0\,, \quad \text{where} \quad Q(\rho_H, \rho_L) \;=\; \omega_2\beta_0^2 - \omega_1\beta_0 + \omega_0$$

Next, we want to demonstrate that $Q(\rho_H, \rho_L)$ is convex in (ρ_H, ρ_L). Calculate

$$\frac{\partial Q}{\partial \rho_H} = \mu_0 \cdot (2\omega_2\beta_0 - \omega_1) - \left(\frac{\beta_0}{\rho_H}\right)^2 \frac{k}{1-\mu_0} + \frac{k\,\mu_0}{1-\mu_0} \quad \Rightarrow$$

$$\frac{\partial^2 Q}{\partial \rho_H^2} = 2\omega_2\mu_0^2 - 2\frac{k\,\mu_0}{1-\mu_0}\frac{\beta_0}{\rho_H^2} + 2\frac{k\beta_0\rho_L}{\rho_H^3}$$

Substitute β_0 and ω_2 above to obtain

$$\frac{\partial^2 Q}{\partial \rho_H^2} = 2\omega_2\mu_0^2 - 2\frac{k\,\mu_0^2}{1-\mu_0}\frac{1}{\rho_H} + 2k(1-\mu_0)\frac{\rho_L^2}{\rho_H^3} = 2\mu_0^2 + 2k(1-\mu_0)\frac{\rho_L^2}{\rho_H^3}\,.$$

Similarly, calculate the partial derivative with respect to ρ_L and subsequently substitute ω_2 to obtain

$$\frac{\partial Q}{\partial \rho_L} = (1-\mu_0)(2\omega_2\beta_0 - \omega_1) = -(1-\mu_0)\omega_1 + 2(1-\mu_0)\beta_0 + 2k\left(\mu_0 + (1-\mu_0)\frac{\rho_L}{\rho_H}\right)\,.$$

Differentiate the first expression above with respect to ρ_L, which appears only in β_0, to obtain the second order partial derivative below. Also, to obtain the cross-derivative below differentiate the second equivalent expression above with respect to ρ_H.

$$\frac{\partial^2 Q}{\partial \rho_L^2} = (1-\mu_0)^2 2\omega_2\,, \qquad \text{and} \qquad \frac{\partial^2 Q}{\partial \rho_L \partial \rho_H} = 2(1-\mu_0)\left(\mu_0 - \frac{k\rho_L}{\rho_H^2}\right)\,.$$

As a result, both of the second order partial derivatives are positive. Hence for $Q(\rho_H, \rho_L)$ to be convex it suffices that

$$\frac{\partial^2 Q}{\partial \rho_H^2}\frac{\partial^2 Q}{\partial \rho_L^2} \geq \left(\frac{\partial^2 Q}{\partial \rho_L \partial \rho_H}\right)^2 \quad \Leftrightarrow$$

$$4(1-\mu_0)^2\omega_2\left(\mu_0^2 + k(1-\mu_0)\frac{\rho_L^2}{\rho_H^3}\right) \geq 4(1-\mu_0^2)^2\left(\mu_0^2 - 2\mu_0\frac{k\rho_L}{\rho_H^2} + \frac{k^2\rho_L^2}{\rho_H^4}\right)$$

Cancel out the $4(1-\mu_0)^2$ terms and substitute ω_2 to equivalently obtain

$$\mu_0^2 + k(1-\mu_0)\frac{\rho_L^2}{\rho_H^3} + \frac{k\mu_0^2}{1-\mu_0}\frac{1}{\rho_H} + \frac{k^2\rho_L^2}{\rho_H^4} \geq \mu_0^2 - 2\mu_0\frac{k\rho_L}{\rho_H^2} + k^2\frac{\rho_L^2}{\rho_H^4} \quad \Leftrightarrow$$

$$(1-\mu_0)\frac{\rho_L^2}{\rho_H^2} + \frac{\mu_0^2}{1-\mu_0} + 2\mu_0\frac{\rho_L}{\rho_H} \geq 0\,,$$

which holds. Hence we have demonstrated that $Q(\rho_H, \rho_L)$ is convex, for all (ρ_H, ρ_L) such that $\rho_L < \theta_L/\theta_H < \rho_H$, which implies that the set of (ρ_H, ρ_L) for which $Q(\rho_H, \rho_L) < 0$ is convex. Therefore, the set of (ρ_H, ρ_L) for which $\mu_0 > \frac{\hat{\beta}-\rho_L}{\rho_H-\rho_L}$ is convex.

Next, consider the linear constrain $\rho_L < \theta_L/\theta_H$. In particular, note that

$$\rho_L \;=\; \frac{\theta_L}{\theta_H} \quad\Rightarrow\quad \beta_0 \;\geq\; \frac{\theta_L}{\theta_H}$$

But the tangency point $\hat{\mu}$ is always in $[0, \mu^*]$, which implies that $\hat{\beta} \leq \theta_L/\theta_H$. Then $\beta_0 \geq \hat{\beta}$, which in turn implies that

$$Q\Big(\rho_H, \frac{\theta_L}{\theta_H}\Big) \;\leq\; 0$$

As a result, $Q(\rho_H, \theta_L/\theta_H) \leq 0$ for all $\rho_H \leq \hat{\rho}_H = \min\Big\{1\,,\, \frac{\theta_L/\theta_H}{\mu_0}\Big\}$, which is our second linear constrain. This describes the boundary of the set on the vertical axis that keeps ρ_L constant.

Next, suppose that $\mu_0 \geq \theta_L/\theta_H$, then

$$\mu_0\hat{\rho}_H + (1-\mu_0)\rho_L \;\geq\; \mu_0\hat{\rho}_H\,\frac{\theta_L}{\theta_H} \;\geq\; \hat{\beta}_0\,,$$

as a result $Q(\hat{\rho}_H, \rho_L) \leq 0$ for all $\rho_L \leq \theta_L/\theta_H$. Suppose instead that $\mu_0 < \theta_L/\theta_H$, and let $\tilde{\rho}_L$ denote the unique solution of ρ_L for $Q(1, \rho_L) = 0$. Then again $Q(1, \tilde{\rho}_L) = 0$. To gather the above result let $\hat{\rho}_L = 0$ in the first case, and $\hat{\rho}_L = \tilde{\rho}_L$ in the second, so that

$$(\hat{\rho}_L, \hat{\rho}_H) \;=\; \begin{cases} (\tilde{\rho}_L, 1) & \text{, if } \mu_0 < \theta_L/\theta_H \\ \Big(0, \frac{\theta_L/\theta_H}{\mu_0}\Big) & \text{, if } \mu_0 > \theta_L/\theta_H \end{cases}$$

Then the horizontal line that connects $(\hat{\rho}_L, \hat{\rho}_H)$ with $(\theta_L/\theta_H, \hat{\rho}_H)$ is the northern boundary of the points for which information provision is optimal, and the vertical line that connects $(\theta_L/\theta_H, \theta_L/\theta_H)$ with $(\theta_L/\theta_H, 1)$ is the eastward. Finally, the set is convex as $Q(\rho_H, \rho_L)$ is convex, and its bounded on the left by $\hat{\rho}_L$ and below by θ_L/θ_H.

To result on the comparison between the first and second best follows from noting that Proposition 1.2, on the online appendix, implies that $\hat{\beta}_i$ is increasing in Ψ_i, the relevant value of which for the first best in zero and for the second best $\mu_0\rho_H$. □

As in the baseline model, a case of special interest is when the buyer's type is the same under both sellers.

Corollary 4. *Suppose that the buyer's type is perfectly correlated across sellers, that is $\rho_L = 0$ and $\rho_H = 1$, then no information provision is optimal.*

***Proof*.** First, suppose that $\rho_H\mu_0 \geq \theta_L/\theta_H$, which for $\rho_H = 1$ becomes $\mu_0 \geq \theta_L/\theta_H$. Then we have already argued that no information provision is optimal. Suppose instead that

$\rho_H \mu_0 < \theta_L/\theta_H$, which for $\rho_H = 1$ becomes $\mu_0 < \theta_L/\theta_H$. Then $J(\mu_0) = 0$ has at most two solutions since it changes monotonicity at most once. One of the solutions is μ^*, and it is ease to see that the other is μ_0. But then it has to be that $\mu_0 < \mu^*$, since J always falls to the point $(\mu^*, 0)$ from above. Hence, μ_0 is below the maximum of J, which implies that $\mu_0 < \hat{\mu}$. Thus, $J(\mu_0) = \mathcal{J}(\mu_0)$, which implies that no information provision is optimal. □

References

Battaglini, M. (2005), 'Long-term contracting with markovian consumers', *The American economic review* **95**(3), 637–658.

Calzolari, G. and Pavan, A. (2006), 'On the optimality of privacy in sequential contracting', *Journal of Economic theory* **130**(1), 168–204.

Calzolari, G. and Pavan, A. (2008), 'On the use of menus in sequential common agency', *Games and Economic Behavior* **64**(1), 329–334.

Calzolari, G. and Pavan, A. (2009), 'Sequential contracting with multiple principals', *Journal of Economic Theory* **144**(2), 503–531.

Dworczak, P. (2016*a*), 'Mechanism design with aftermarkets: Cutoff mechanisms.'.

Dworczak, P. (2016*b*), 'Mechanism design with aftermarkets: On the optimality of cutoff mechanisms.'.

Ely, J. C. (2017), 'Beeps', *The American Economic Review* **107**(1), 31–53.

Eső, P. and Szentes, B. (2017), 'Dynamic contracting: An irrelevance result', *Theoretical Economics* (12), 109–139.

Garrett, D. F. and Pavan, A. (2012), 'Managerial turnover in a changing world', *Journal of Political Economy* **120**(5), 879–925.

Gentzkow, M. and Kamenica, E. (2011), 'Bayesian persuasion', *American Economic Review* **101**(6), 2590–2615.

Inostroza, N. and Pavan, A. (2017), 'Persuasion in global games with application to stress testing', *Economist* .

Pavan, A., Segal, I. and Toikka, J. (2014), ‘Dynamic mechanism design: A myersonian approach’, *Econometrica* **82**(2), 601–653.

Roesler, A.-K. and Szentes, B. (2017), ‘Buyer-optimal learning and monopoly pricing’, *forthcoming American Economic Review* .