Screening with behavioral buyers*
PRELIMINARY AND INCOMPLETE

Nicolas Pastrian†

May 12, 2021

Abstract

This paper studies a monopolist screening problem where some buyers are *behavioral*: they fail to recognize the incentive scheme and report truthfully regardless of the mechanism implemented. We characterize the optimal mechanism in this setting and provide a interpretation of the proposed solution in the context of pricing. We present three extensions to the basic model. First, we analyze the impact of considering ex-ante participation constraints instead of interim participation, and show that the designer could use this slackness to better discriminate between behavioral and non-behavioral buyers. In the second extension, we analyze the case of costly sophistication, where behavioral buyers could imitate strategic buyers paying a fee. We show that this possibility limits the power of the seller to exploit behavioral buyers. Finally, we explore the use of constrained mechanisms in which the seller is constrained on the number of different contracts he could offer to the buyers. For this setting we show that the solution involves moving from the optimal mechanism which ignores behavioral buyers to the mechanism that ignores incentive compatibility constraints but allows some types to deviate.

*I thank Luca Rigotti and Richard Van Weelden for providing guidance in this project. I also thank Sofia Moroni, Svetlana Kosterina, Arie Beresteanu, Benjamin Matta, Mark Azic, and participants of the Microeconomic Theory Brownbag at University of Pittsburgh for useful comments.

†Department of Economics, University of Pittsburgh. Email: nip59@pitt.edu

1 Introduction

Mechanism design has studied how to design optimal allocation rules, addressing that agents will respond to the incentives generated by the rules themselves. However, designing new mechanisms and implementing reforms to existing economic systems could have unintended consequences on the behavior of the economic agents interacting with them. For example, it is possible that some agents fail to recognize the incentives generated by a new scheme, deviating from the behavior we should expect from them in equilibrium. While an alternative interpretation to equilibrium behavior is it would arise as a response to repeated interactions, even under this interpretation deviations to traditional equilibrium behavior could survive if interactions have low frequency, or when completely novel interactions are studied. For example, college and school admission problems, a common application of mechanism design techniques, are problems which families usually only face a limited number of times during lifetime. In consequence, while the idea of equilibrium behavior makes sense for a well known environment, the introduction of changes on the system could make equilibrium behavior not as obvious as the classic theory suggests. This unexpected behavior could have consequences in terms of both the final allocation implemented and the revenue generated by the new implemented mechanism. Hence, in the design of such reforms it could be necessary that the designer take into account the non-equilibrium behavior to either improve the outcome obtained, or guarantee that the desired goal is satisfied.

Here, we study a simple screening problem where the agent could fail to recognize the incentive scheme and report truthfully regardless of the mechanism implemented. We characterize the optimal mechanism in this setting and study some extensions, including an environment with costly sophistication.

Our setting is a multidimensional environment where buyers types are characterized by their valuation (either low or high) and their informational type (either "behavioral" or "strategic"[1]). In order to limit the complexity which naturally arise on multidimensional environments, we focus on a simple monopolistic screening model where the seller's objective is to maximize his expected profits. We present some extensions that accounts for ex-ante participation constraints, information acquisition and restricted messages spaces.

[1]We will refer to agents not perfectly responding to the incentive schemes as behavioral, however we think of this agents as being unaware of the incentives created by the mechanism, and hence relying on an alternative strategy which we will assume is always truthful reporting. In the same way, strategic players will be the ones that react optimally to the mechanism implemented because they are aware of the incentives generated by such mechanism.

In our baseline screening model the seller exploits behavioral types with high valuation but is unable to exploit low valuation agents. In one of our extensions, where we allow for small variations in the environment, the seller will be able to further separate behavioral and strategic types, fully exploiting behavioral ones.

Instead, if we restrict the size of the menu that the seller could offer then the solution reduces to either offer the first best mechanism but ignore that some buyers will deviate from their designed contract, or implementing the solution to the standard screening problem ignoring the presence of behavioral players.

For our information acquisition variation, we show that if the behavioral type could learn about his own restrictions and behave as strategic agents then basically the seller is forced to offer the standard contract unless the cost of learning is too big.

The remaining of the paper is organized as follows. Section 2 offers a review of the related literature. Section 3 presents the main model, while Section 4 presents some benchmarks and the mechanism proposed as solution. The optimality of this mechanism is proved in Section 5. Some extensions are presented in Section 6. Finally, Section 7 concludes.

2 Related literature

A recent paper is closely related to the model studied here. Li and Peters (2020) study an auction environment where the mechanism is not observed by some players. They refer to this players as uninformed and analyze all equilibria that could arise in this context, without appealing to a revelation principle directly. They show that there are both communicative and uncommunicative equilibria in their model, but that under some conditions the equilibrium is essentially uncommunicative for uninformed players. Here instead we use a revelation principle argument to simplify the analysis and focus only on partial implementation, i.e., equilibrium selection by the seller.

Since in our model the information of the buyer has two components, this paper is related to multidimensional mechanism design. However, here we focus on a model where only one of such pieces is payoff relevant, as opposed to a setting where such both pieces represent the value associated with different goods as in Manelli and Vincent (2007), or externalities as in Jehiel et al. (1999). As noted by Börgers (2015), the analysis of such environments could become cumbersome due to envelope and weak monotonicity conditions being only necessary but no longer sufficient for implementation. Moreover, in our modeling, the cyclic monotonicity used in Rochet (1987) becomes only sufficient but no longer necessary due to a relaxation of some constraints. To overcome such difficulties, we focus on a simple revenue (profit) maximization environment where characterization becomes less cumbersome.

Salant and Siegel (2018) studies a contract design problem where the principal could choose the frame that will determine the agent valuations for each potential contract. However, this model limits to the selection of a single frame. Instead, Ostrizek and Shishkin (2019) studies a model of monopolistic screening where the buyer is sophisticated but dynamic inconsistent, and the seller could design both the contract and which frames the buyer would face in each stage, by designing the complete extensive form game the buyer will face. In our model, there is no frames and seller limits to design a menu of contracts, that is, a menu of prices and quality provided, but doesn't have any further impact on buyer's valuation.

Severinov and Deneckere (2006) studies a continuous version of the basic model presented here and notes that there is no exclusion in presence of honest types. However, in our model exclusion could remain optimal in presence of honest types. In a related work, Saran (2011) studies a bilateral trade model in presence of behavioral types. He focuses on efficiency and shows that by fully extracting the rent from naive traders efficiency could be improved. He also shows that for a fixed mechanism the impact of behavioral players on efficiency is ambiguous.

There are some papers that study behavioral mechanism design. An earlier example is the work of Eliaz (2002). There the author studies implementation where at most k players behavior could not be predicted and shows than a stronger form of incentive compatibility must be satisfied for implementation. Recently, Clippel et al. (2018) studies implementation under the k-level reasoning model. They characterize the conditions required for a mechanism to be implementable in this context.

The presence of behavioral types in our model could be interpreted as a robustness requirement. However, this is a different approach from the one presented in the classic work of Bergemann and Morris (2005) where the robustness component comes from the specification of the beliefs that player have over other players information. That leads to dominance of using dominant strategy mechanism in some environments. An alternative approach was explored in Li (2017) where robustness requires not only be dominant strategy mechanisms but also of a particular class: where strategies remain optimal even if we consider only "obvious" deviations. However, both approaches ignore any further behavioral consideration. Several other robustness approaches are explored in a recent survey by Carroll (2019).

There is also some experimental evidence on the failure of dominant strategy mechanisms. For example, Rees-Jones and Skowronek (2018) shows there is untruthful reporting in the medical residency matching allocation.

This paper is also slightly related to literature of learning in mechanism design settings, like Bobkova (2019) and Gleyze and Pernoud (2020). However, this work differs from the previous papers because here there is no learning about valuations, neither about the own player nor about the other players. Instead, learning is studied in terms of the sophistication level and the details of the mechanism itself.

3 Model

Consider a seller which produces a single good of quality $q \geq 0$ at cost $c(q) = q^2/2$. There is a continuum of buyers, each one with privately observed valuation $\theta \in \{\theta_L, \theta_H\}$, where $\theta_H > \theta_L > 0$, and $\Pr(\theta = \theta_H) = \alpha \in (0,1)$ is the proportion of buyers with high valuation. Alternatively, we could assume that there is a single buyer which draws a high valuation with probability α.

We will assume that buyers have quasilinear preferences given by

$$\theta q - t$$

where $t \in \mathbb{R}$ is the transfer paid to the seller. The seller's objective is to maximize his expected profits.

We depart from the traditional setting by assuming that a fraction of buyers fails to recognize the incentives generated by the mechanism. We could interpret this as some buyers facing infinite lying cost, being intrinsically honest, or being uninformed about the details of the mechanism. We will refer to this type of buyers as *behavioral*. If a player is not behavioral, then he is completely strategic, faces no cost from lying and is able to choose his contract or message in an optimal way. Behavioral and strategic types will be denoted by B and S respectively. Notice that in this environment the type of a buyer has two dimensions: it contains both his valuation and whether he is behavioral or not. We will denote the fraction of strategic types by $\gamma \in (0,1)$. Before introducing the details of the behavior of the different types of buyers, we will define what is a mechanism in this context.

We will restrict attention to direct mechanisms: $\Gamma = (q_i^S, t_i^S, q_i^B, t_i^B)_{i\in\{L,H\}}$, where q_i^j and t_i^j are the quality received and the price paid by a buyer which reports being of type (θ_i, j) with $j \in \{S, B\}$. We will refer to the pair (q_i^j, t_i^j) as a contract.

We will assume that a behavioral agent in a direct mechanism always report his true type (both his valuation and his informational status). Instead, a strategic buyer will be able to report any message available and will report truthfully only if its optimal for him doing so. While the assumption over the actions of the behavioral buyers seems restrictive we justify its introduction for two reasons. First, it allows to simplify the characterization of the behavior of agents that doesn't completely react to the incentives provided by the mechanism. Second, this assumption follows from a revelation principle argument on the behavior of this agents, and it is without loss of generality as long two conditions are satisfied with respect to the behavioral buyers: (i) that indeed their behavior doesn't depend on the mechanism offered, and (ii) that their behavior is rich enough so that the seller could identify their valuations[2].

Moreover, notice that the actual implementation of this class of mechanisms doesn't require behavioral players reporting directly whether they are strategic or not: a designer could include as valid messages the valuations and ask honest players report only such, while use "special" messages for strategic players, including an extra piece of information or using a completely different set of messages.

Hence, by this revelation principle argument, both the use of direct mechanisms and the restriction on the behavior of non-strategic buyers is without loss in the current setting[3].

We will denote by $R(\Gamma)$ the expected profits of the seller for a mechanism Γ in which each type report truthfully, i.e.,

$$\begin{aligned} R(\Gamma) = \ & \gamma\left[\alpha\left(t_H^S - c(q_H^S)\right) + (1-\alpha)\left(t_L^S - c(q_L^S)\right)\right] \\ & +(1-\gamma)\left[\alpha\left(t_H^B - c(q_H^B)\right) + (1-\alpha)\left(t_L^B - c(q_L^B)\right)\right] \end{aligned}$$

In order to $R(\Gamma)$ to represent the actual expected profits received by the seller it is necessary that agents comply with their contract, i.e., that each type actually chooses the contract designed for his true type. Notice that this is never a concern with respect to behavioral types since they will always report truthfully by assumption. In contrast, for strategic type some kind of incentive compatibility constraint should be satisfied in order to guarantee their compliance. In particular, the traditional incentive compatibility constraint are necessary for this types, i.e., that (strategic) low types don't want to imitate (strategic) high types, and that (strategic) high types don't want to imitate (strategic) low types. However, new incentive compatibility constraint arise due to the possibility that strategic types could also imitate behavioral types. Hence, the set of incentive compatibility constraints required for the strate-

[2] Obviously, this last condition could be relaxed if we allow for correlation between the valuation and the informational types.

[3] A formal proof of this argument is available upon request.

gic types is larger than in the standard mechanism design problem without behavioral buyers.

We formally introduce the incentive compatibility constraints that a mechanism need to satisfy in this context in the next definition.

Definition 1. *We say that a mechanism is* incentive compatible *if it satisfies the following conditions*

$$\theta_H q_H^S - t_H^S \geq \theta_H q_H^B - t_H^B \tag{$IC_{HS,HB}$}$$

$$\theta_H q_H^S - t_H^S \geq \theta_H q_L^S - t_L^S \tag{$IC_{HS,LS}$}$$

$$\theta_H q_H^S - t_H^S \geq \theta_H q_L^B - t_L^B \tag{$IC_{HS,LB}$}$$

$$\theta_L q_L^S - t_L^S \geq \theta_L q_H^B - t_H^B \tag{$IC_{LS,HB}$}$$

$$\theta_L q_L^S - t_L^S \geq \theta_L q_H^S - t_H^S \tag{$IC_{LS,HS}$}$$

$$\theta_L q_L^S - t_L^S \geq \theta_L q_L^B - t_L^B \tag{$IC_{LS,LB}$}$$

For most of the paper we will consider only interim individually rational mechanisms, this is required to bound the potential revenue that the designer could get from a behavioral agent[4]. Introducing this type of constraints also rule out the possibility of offering an extremely adversarial contract to behavioral types in order to make such contracts unattractive for strategic players (which would essentially reduce the problem to the standard environment without behavioral buyers).

Notice also that such participation constraints are only required for the behavioral types since incentive compatibility will automatically imply that they should also hold for strategic types. In other words, the individual rationality constraints for strategic agents are always redundant. In consequence, we will completely omit such constraints in the characterization below. Hence, the relevant participation constraints in our setting will be

$$\theta_H q_H^B - t_H^B \geq 0 \tag{IR_{HB}}$$

$$\theta_L q_L^B - t_L^B \geq 0 \tag{IR_{LB}}$$

[4]Ex-ante participation constraints are explored in Section 6 below.

We will consider ex-post[5] participation constraints here since they could be interpreted as the possibility of rejecting the contract after it is implemented.[6]

We consider a mechanism being feasible if it satisfy all IC and IR constraints above.

Definition 2. *We say that a mechanism is* feasible *if it is incentive compatible, and satisfies* IR_{HB} *and* IR_{LB}.

Given the above description, we can write down the seller's problem as

$$\max_{\Gamma} R(\Gamma)$$

subject to Γ being a feasible mechanism.

4 Benchmarks and the optimal mechanism

There are two natural benchmarks for the basic model. First, the standard problem without behavioral types. Second, the optimal mechanism which ignores the incentive compatibility constraints. We will refer to the solution to this problems as the *standard mechanism without behavioral buyers* and the *full information mechanism* respectively.

It is well known, and easy to verify, that the optimal mechanism for the problem without behavioral types is characterized by

$$q_L = \max\left\{\frac{\theta_L - \alpha\theta_H}{1-\alpha}, 0\right\} \text{[7]}$$

$$q_H = \theta_H$$

$$t_L = \theta_L q_L$$

$$t_H = \theta_H q_H - (\theta_H - \theta_L) q_L$$

[5] Notice that interim and ex-post requirements coincide in our setting.

[6] However, this could rule out the possibility of using some of the rents extracted from the behavioral types in order to implement different allocations, as Saran (2011) have shown for the bilateral trade model. Hence, this is *not* without loss of generality. We will explore the impact of ex-ante participation constraints in Section 6.

[7] Notice that this is equivalent to

$$\max\left\{\theta_L - \frac{\alpha}{1-\alpha}\left(\theta_H - \theta_L\right), 0\right\}$$

Hence, we will observe exclusion of the buyers with low valuation if $\theta_L \leq \alpha\theta_H$ (or $\theta_L < \frac{\alpha}{1-\alpha}(\theta_H - \theta_L)$). We will refer to this solution as the *standard mechanism without behavioral buyers*.

For the second benchmark, we consider the case in which the seller could perfectly identify the valuation of each buyer. Hence, the optimal mechanism in this case is given by $t_i = \theta_i q_i$ and $q_i = \theta_i$ for $i = L, H$. Notice that this involves providing the efficient quality to each buyer and extract all the surplus from them. We will refer to this solution as the *full information mechanism*.

Two remarks could be made with respect to the above mechanisms. First, the standard mechanism without behavioral buyers remains feasible in our environment without further modifications. However, we will show that it is possible to improve over the profits generated by such mechanism in our setting. In contrast, the full information mechanism could not be properly implemented in our setting. While it remains possible to implement it for behavioral types, the presence strategic buyers makes them deviate to the contract offered to behavioral types, violating incentive compatibility, and making implementation infeasible.

Now we proceed to characterize the main result for the basic model. We characterize the solution to the seller's problem below. We will refer to its solution as the *optimal mechanism*.

Proposition 1. *The optimal mechanism is given by*

$$q_L^S = \max\{\theta_L - \frac{\gamma\alpha}{1-\alpha}\left(\theta_H - \theta_L\right), 0\}$$

$$q_H^S = \theta_H$$

$$q_L^B = \max\{\theta_L - \frac{\gamma\alpha}{1-\alpha}\left(\theta_H - \theta_L\right), 0\}$$

$$q_H^B = \theta_H$$

$$t_L^S = \theta_L q_L^S$$

$$t_H^S = \theta_H q_H^S - (\theta_H - \theta_L) q_L^B$$

$$t_L^B = \theta_L q_L^B$$

$$t_H^B = \theta_H q_H^B$$

Under the mechanism proposed above all buyers with high valuation receive the efficient quality provision but the transfers they need to pay will depend on whether they are strategic or behavioral. Instead, for low valuation buyers a single contract is offered and the quality provided in this contract is always below the efficient level. Moreover, there is exclusion of the buyers with low valuation if $\theta_L < \frac{\gamma\alpha}{1-\alpha}(\theta_H - \theta_L)$. However, exclusion will be observed less often since this requires a lower threshold in comparison to the case where behavioral buyers are absent. Finally, note that under this mechanism only strategic buyers with high valuation get informational rents.

While the mechanism proposed could seems unpractical, we will see that in the context of a price setting environment the proposed mechanism has a natural interpretation. The seller will offer two different qualities for the product: a cheap version which is provided an inefficiently low quality, and a premium version of the product, provided at the efficient quality level for high valuation buyers. For the cheap version a single price is offered to all buyers (no discount or rebate is available), while for the premium version two different prices are offered: a discounted price for "sophisticated" (strategic) buyers and the normal price for "simple" (behavioral) buyers. Here, in order to get access to this discounted price the buyer would need to fulfill some requirement (bring a coupon, ask explicitly for a discount, use a "price match", check a box, fill a form, etc) which is only recognized by the buyers informed of such possibility. A similar interpretation could also be given in an income tax setting, where only some citizens claim returns available for them.

The following results follows directly from Proposition 1. They offer some comparative statics in terms of the fraction of strategic buyers γ.

Corollary 1. *Fix θ_L, θ_H and α. There exists $\bar{\gamma}$ such that for $\gamma < \bar{\gamma}$ there is no exclusion.*

In particular, $\overline{\gamma}$ is defined as

$$\overline{\gamma} = \sup \left\{ \gamma \in [0,1] \,\middle|\, \theta_L - \frac{\gamma\alpha}{1-\alpha}(\theta_H - \theta_L) \geq 0 \right\}$$

and we have that $\overline{\gamma} = \left(\frac{1-\alpha}{\alpha}\right) \frac{\theta_L}{\theta_H - \theta_L}$ if $\alpha\theta_H \geq \theta_L$, and $\overline{\gamma} = 1$ otherwise. Notice also that the optimal mechanism converges to the full information mechanism as γ approaches zero, and it converges to the standard mechanism without behavioral buyers as γ approaches one. Hence, the optimal mechanism presented here satisfy a natural notion of continuity with respect to γ.

The result above also implies that exclusion is less likely in this setting. Moreover, if there are enough behavioral buyers, we can always get no exclusion in the optimal mechanism.

We have that the welfare, defined as the sum of the utility of the buyers and the seller's profits, is decreasing and concave in γ (strictly concave for $\gamma < \overline{\gamma}$ and constant afterwards), hence a higher proportion of behavioral buyers increases efficiency. Similarly, the seller's profits are decreasing and convex in γ (strictly for $\gamma < \overline{\gamma}$ and constant afterwards), meaning that the presence of behavioral types makes the seller better off.

These main comparative static with respect to the fraction of strategic buyers are summarized in the following corollary.

Corollary 2. *Both welfare and seller's profits are increasing in the fraction of behavioral buyers (decreasing in γ).*

5 Proof of Proposition 1

In this section we present the proof of the main result (Proposition 1). The procedure will be carried over in several steps, the first one being characterize the payment scheme for the behavioral types. In particular, we show that all the rents of the behavioral types will be always extracted, although their allocation could be potentially distorted.

Lemma 1. *Let Γ be a profit maximizing mechanism. Then, $t_i^B = \theta_i q_i^B$ for $i = L, H$.*

Proof. Suppose not, that is Γ is a revenue maximizing mechanism but $t_i^B < \theta_i q_i^B$, i.e., the participation constraint for type (θ_i, B) is not binding.

Consider $\hat{\Gamma}$ such that $\hat{\Gamma}_{-\hat{t}_i^B} = \Gamma_{-t_i^B}$ and $\hat{t}_i^B = t_i^B + \epsilon$, with $\epsilon > 0$ and $\hat{t}_i^B \in (t_i^B, \theta_i q_i^B)$.

Since (θ_i, B) has strictly positive probability (which follows from $\gamma > 0$ and $\alpha \in (0,1)$), we have that $R(\hat{\Gamma}) > R(\Gamma)$. Clearly, $\hat{\Gamma}$ is a feasible mechanism (satisfies IC and IR constraints), which implies Γ is not a profit maximizing mechanism.

Hence, we must have $t_i^B = \theta_i q_i^B$ in a profit maximizing mechanism. $\square$

This result help us to simplify the seller's problem, allowing us to discard some constraints and reducing the number of choice variables.

Claim 1. *$IC_{HS,HB}$ and $IC_{LS,HB}$ are redundant.*

Proof. Replacing the above results in those constraints we get

$$\theta_H q_H^S - t_H^S \geq \theta_H q_H^B - \theta_H q_H^B = 0$$

$$\theta_L q_L^S - t_L^S \geq \theta_L q_H^B - \theta_H q_H^B = -(\theta_H - \theta_L) q_H^B \leq 0$$

also, from $IC_{HS,LB}$ and $IC_{LS,LB}$,

$$\theta_H q_H^S - t_H^S \geq (\theta_H - \theta_L) q_L^B \geq 0$$

$$\theta_L q_L^S - t_L^S \geq 0$$

□

The next results will be useful to show that the mechanism in Proposition 1 is indeed optimal.

Claim 2. *$q_L^B \leq \theta_L$ in a profit maximizing contract.*

Proof. If q_L^B is above the efficient level ($q_L^B > \theta_L$), then both $IC_{HS,LB}$ and profits are decreasing in q_L^B. Note that $IC_{HS,LB}$ is the only condition where q_L^B enters. Hence, by reducing q_L^B the designer increases his revenue without changing incentives since incentive compatibility constraints are easier to satisfy. □

Claim 3. *In a profit maximizing mechanism at least one of $IC_{HS,LS}$ or $IC_{HS,LB}$ is binding.*

Proof. Let Γ be a solution to the seller's problem. Suppose both $IC_{HS,LS}$ and $IC_{HS,LB}$ are not binding for this mechanism.

Then there exist a contract $\hat{\Gamma}$ such that (i) $\hat{\Gamma}_{-\hat{t}_H^S} = \Gamma_{-t_H^S}$; (ii) $\hat{t}_H^S > t_H^S$; and, (iii) no constraint is violated. The last statement comes from the fact that the only conditions that are decreasing in t_H^S are $IC_{HS,LS}$ and $IC_{HS,LB}$, but they are not binding by assumption.

However, $R(\hat{\Gamma}) > R(\Gamma)$. Thus, Γ can't be a profit maximizing mechanism.

□

Now, we are in position to prove the main result.

Proof of Proposition 1. From Claim 3, it is without loss to look at mechanisms where either $IC_{HS,LS}$ or $IC_{HS,LB}$ is binding.

Suppose $IC_{HS,LB}$ is binding. Then, by Claim 1,

$$t_H^S = \theta_H q_H^S - (\theta_H - \theta_L) q_L^B$$

Now, using Claim 1 again and rearranging the remaining constraints we get

$$(\theta_H - \theta_L) q_L^B \geq \theta_H q_L^S - t_L^S \tag{1}$$

$$\theta_L q_L^S - t_L^S \geq 0 \tag{2}$$

$$\theta_L q_L^S - t_L^S \geq (\theta_H - \theta_L)\left(q_L^B - q_H^S\right) \tag{3}$$

We can rearrange above equations to obtain necessary conditions for the remaining allocations. In particular, from (1) and (2),

$$\theta_L q_L^S \geq t_L^S \geq \theta_H q_L^S - (\theta_H - \theta_L) q_L^B$$

which implies

$$q_L^B \geq q_L^S.$$

Similarly, from (1) and (3),

$$\theta_L q_L^S - (\theta_H - \theta_L)\left(q_L^B - q_H^S\right) \geq t_L^S \geq \theta_H q_L^S - (\theta_H - \theta_L) q_L^B$$

which give us,

$$q_H^S \geq q_L^S$$

Using the results above, we can rewrite the seller's profits as

$$\gamma\left[\alpha\left(\theta_H q_H^S - (\theta_H - \theta_L) q_L^B - \frac{q_H^2}{2}\right) + (1-\alpha)\left(t_L^S - \frac{(q_L^S)^2}{2}\right)\right] + (1-\gamma)\left[\alpha\frac{\theta_H^2}{2} + \left(\theta_L q_L^B - \frac{(q_L^B)^2}{2}\right)\right]$$

Moreover, notice that both the revenue and (3) are increasing in q_H^S if $q_H^S \leq \theta_H$, thus setting q_H^S at the efficient level (i.e., $q_H^S = \theta_H$) is indeed optimal.

Since we already have shown that $q_L^B \leq \theta_L$, $q_H^S = \theta_H$ implies (3) is redundant.

From the two remaining constraints, we have that (2) must be binding since both profits and (1) are increasing in t_L^S while (2) is decreasing in t_L^S. Hence, $t_L^S = \theta_L q_L^S$.

Replacing this into the remaining constraint reduces to a monotonicity condition over the quality levels provided to buyers with low valuation

$$q_L^B \geq q_L^S \tag{4}$$

Thus, the seller's problem is reduced to

$$\begin{aligned} \max_{q_L^B, q_L^S} \quad & \gamma\left[\alpha\left(\frac{\theta_H^2}{2} - (\theta_H - \theta_L) q_L^B\right) + (1-\alpha)\left(\theta_L q_L^S - c(q_L^S)\right)\right] \\ & + (1-\gamma)\left[\alpha\left(\frac{\theta_H^2}{2}\right) + (1-\alpha)\left(\theta_L q_L^B - c(q_L^B)\right)\right] \end{aligned}$$

subject to the monotonicity constraint in (4) above.

Consider now the relaxed problem, that is, the problem above ignoring the monotonicity constraint. The first order conditions to this relaxed problem are given by

$$-\gamma\alpha(\theta_H - \theta_L) + (1-\gamma)(1-\alpha)(\theta_L - q_L^S) \leq 0$$

$$\gamma(1-\alpha)(\theta_L - q_L^B) \leq 0$$

From here, it is clear that the solution of the relaxed problem involves

$$q_L^S = \theta_L > q_L^B$$

which violates the monotonicity constraint. Hence, the solution to the original problem must satisfy $q_L^B = q_L^S$. Using this result, the seller's problem is further reduced to

$$\max_{q_L^S} \gamma\alpha\left(\frac{\theta_H^2}{2} - (\theta_H - \theta_L)q_L^S\right) + (1-\alpha)\left(\theta_L q_L^S - c(q_L^S)\right) + (1-\gamma)\alpha\frac{\theta_H^2}{2}$$

the first order condition of this problem is

$$-\gamma\alpha(\theta_H - \theta_L) + (1-\alpha)(\theta_L - q_L^S) \leq 0$$

which implies

$$q_L^S = \max\left\{\theta_L - \frac{\gamma\alpha}{1-\alpha}(\theta_H - \theta_L), 0\right\}$$

From which the mechanism proposed in Proposition 1 is obtained as the solution.

It remains to show that indeed $IC_{HS,LB}$ should be binding in the solution. Let's suppose it is not the case. Then, by Claim 3, $IC_{HS,LS}$ should be binding. Thus, the transfer required from type (θ_H, S) is given by

$$t_H^S = t_L^S + \theta_H(q_H^S - q_L^S)$$

Using this, we can rewrite the seller's problem as

$$\max_{q_L^S, q_L^B, q_H^S, t_L^S} \gamma\left[\alpha\left(t_L^S + \theta_H(q_H^S - q_L^S) - c(q_H^S)\right) + (1-\alpha)\left(t_L^S - c(q_L^S)\right)\right]$$
$$+ (1-\gamma)\left[\alpha\left(\frac{\theta_H^2}{2}\right) + (1-\alpha)\left(\theta_L q_L^B - c(q_L^B)\right)\right]$$

subject to

$$q_H^S \geq q_L^S$$

$$\theta_L q_L^S - t_L^S \geq 0$$

$$\theta_L q_L^S - t_L^S \geq (\theta_H - \theta_L) q_L^B$$

$$q_L^S, q_L^B, q_H^S \geq 0$$

Notice that if the third constraint is binding, we are back to the previous case. Hence, we can ignore such condition in the following analysis. This implies that the second condition should be binding in the solution (i.e., $t_L^S = \theta_L q_L^S$), otherwise we can construct a new mechanism which dominates the proposed solution in terms of profits.

Hence, we can write the profits of the seller as

$$\gamma \left[\alpha \left(\theta_H q_H^S - (\theta_H - \theta_L) q_L^S - c(q_H^S)\right)(1-\alpha)\left(\theta_L q_L^S - c(q_L^S)\right)\right] \\ +(1-\gamma)\left[\alpha\left(\frac{\theta_H^2}{2}\right) + (1-\alpha)\left(\theta_L q_L^B - c(q_L^B)\right)\right]$$

In the relaxing problem (ignoring the monotonicity requirement for the allocation), the first order conditions are given by

$$\gamma\alpha\left(\theta_H - q_H^S\right) \leq 0$$

$$-\gamma\alpha(\theta_H - \theta_L) + \gamma(1-\alpha)(\theta_L - q_L^S) \leq 0$$

$$(1-\gamma)(1-\alpha)(\theta_L - q_L^B) \leq 0$$

which implies that the optimal mechanism should be

$$q_H^S = \theta_H$$

$$q_L^S = \max\left\{\theta_L - \frac{\alpha}{1-\alpha}(\theta_H - \theta_L), 0\right\}$$

$$q_L^B = \theta_L$$

but such contract violates $IC_{HS,LB}$. Thus, the solution to the seller's problem must have $IC_{HS,LB}$ binding. □

6 Extensions

In this section we independently explore three different extensions to the basic model presented in the sections above. In particular, we look at a model which relaxes the participation constraints, a model where behavioral buyers could acquire some extra information, and an environment where the space of messages is constrained.

6.1 Ex-ante participation

Let's consider an alternative formulation requiring ex-ante participation constraints instead of interim participation as before. In other words, a setting where buyers can only opt out from the mechanism allocation before they learn their own valuation. We will assume that behavioral buyers are partially sophisticated in the sense that they are able to calculate the actual payoff they will receive if they participate in the mechanism.

We start relaxing the participation constraints for all types, that is requiring the ex-ante participation constraints for all types.

Definition 3. *We say that a mechanism satisfies ex-ante participation if*

$$\alpha\theta_H q_H^S + (1-\alpha)\theta_L q_L^S \geq 0 \qquad (EAIR_S)$$

and

$$\alpha\theta_H q_H^B + (1-\alpha)\theta_L q_L^B \geq 0 \qquad (EAIR_B)$$

Proposition 2. *An optimal mechanism with ex-ante participation satisfies*

$$q_i^S = q_i^B = \theta_i$$

$$t_i^S = t_i^B = t_i$$

$$\alpha t_H + (1-\alpha)t_L = \alpha\theta_H^2 + (1-\alpha)\theta_L^2$$

$$\theta_H(\theta_H - \theta_L) \geq t_H - t_L \geq \theta_L(\theta_H - \theta_L)$$

for $i = L, H$. Moreover, an optimal mechanism exists.

Proof. Notice that the equations above are compatible and such system allows multiple solutions. Moreover, all those solutions provide the efficient quality to all types of buyers and fully extract the first best level utility from the buyer in expectation. Since profits collected by these mechanisms coincide with the profits on the full information mechanism (i.e., the mechanism which ignores incentive compatibility constraints), they are indeed profit maximizing. □

The structure of the mechanism could be summarized as follows: we start with a mechanism that is incentive compatible and provides the efficient quality level to both valuation types, then we modify transfers in way that the difference between payments remain the same but the expected payment equalize the expected utility generated by such allocation. This arrangement is always feasible and extract all the rents in expectation.

Here, the existence of behavioral buyers doesn't make any difference since without them the optimal mechanism will look exactly the same. Notice that if instead there were no strategic types, then the family of mechanisms characterized above represents only one form that maximizes profits for the seller. Indeed, in such case a mechanism that provides the efficient quality and fully extract ex-post all the rents from both valuation types is also optimal.

An important remark here is that while this mechanism fully extract in expectation, ex-post, the payoff of buyers with high valuation are strictly positive while the payoff of buyers with low valuation are strictly negative.

6.1.1 Ex-ante participation only for behavioral buyers

Now consider the ex-ante participation only for behavioral buyers but return to the interim participation constraints for strategic buyers. That is, strategic buyers could not just observe the full menu of contracts, but also could opt out of the mechanism after they have looked at them and learned their valuation; while the behavioral buyers must commit to accept their contract beforehand, knowing that they will report truthfully afterwards. In the context of the pricing example, this could be interpret as having behavioral buyers deciding whether to enter or not to the shop but knowing that after entering they will probably return with something.

In this context, the relevant participation constraints for strategic buyers are

$$\theta_H q_H^S - t_H^S \geq 0 \tag{IR_{HS}}$$

$$\theta_L q_L^S - t_L^S \geq 0 \tag{IR_{LS}}$$

This means that strategic buyers not only could identify contracts better than behavioral ones, but that also they doesn't face any self-control problem in this context.

We now formally define the participation requirements in this section

Definition 4. *We say that a mechanism satisfies ex-ante participation only for behavioral types if it satisfies $EAIR_B$, IR_{HS}, and IR_{LS}.*

There are two key inequalities defining the optimal mechanism in this case: that the strategic buyer with high valuation doesn't want to deviate to the contract of the behavioral buyer with high, nor the contract of the behavioral buyer with low valuation.

Using this insight, we can construct a candidate mechanism starting from the optimal mechanism with interim participation.

First, note that the expected utility of a behavioral buyer from the interim mechanism is zero since both valuation types are getting zero utility. However, by increasing the transfer from the behavioral buyer with low valuation and decreasing the transfer from the behavioral buyer with high valuation in a way that the expected payment remains the same, the mechanism remains incentive compatible satisfies the ex-ante participation constraint and generates as much revenue as the interim mechanism.

However, now there is space to increase the provision of quality to the behavioral buyers with low valuation without violating any other constraint. By doing so, the seller is allowed to provide an allocation closer to the first best allocation and extract more rents from a behavioral buyer. This can be done as long as (i) the payoff of the behavioral buyer with high valuation is not increased too much so that the strategic type wants imitate it, and (ii) the payoff that the strategic buyer with high valuation could get from imitating the low behavioral type remains low enough. The last condition comes from the complementarity between valuations and quality: a higher quality provision for the (behavioral) buyers with low valuation could be tempting for the (strategic) buyer with high valuation. This is the reason why the two constrains highlighted before are key to characterize the optimal mechanism in this context.

In particular, we have that for $\theta_L > \frac{\theta_H}{2-\alpha}$ the optimal mechanism fully discriminate between behavioral and strategic types: for behavioral buyers the contracts offered are analogous to the ones used in the case of ex-ante participation, providing the efficient quality to both valuation types used is similar to the one of the previous section which provides the efficient and fully extracting their rents in expectation; while for the strategic buyers the contracts offered are exactly the ones offered if behavioral buyers were absent. As in the case with ex-ante participation for both behavioral and strategic buyers, there is not an unique optimal mechanism since the contracts offered to behavioral buyers are not uniquely defined. The reason why the strategic buyers are "reverted" to the mechanism without behavioral types is that the optimal contracts for them only depends on the relative proportion of low and high types taking such contracts, since behavioral buyers receive different contracts in this case, the proportion of behavioral buyers will have no impact on the contracts offered to strategic buyers. This is contrast to the case with interim participation where both low valuation types receive the same contract.

Having several contracts satisfying the required conditions is only possible for high values of θ_L. For lower values, two cases arise. First, for $\theta_L \in (\gamma\alpha\theta_H, \frac{\theta_H}{2-\alpha})$, the two key inequality constraints defining the mechanisms for the behavioral buyers collapse and both are always binding in the solution. In this case, it is no longer true that the efficient quality is provided to both behavioral types. The second case arise for $\theta_L < \gamma\theta_H$. In this case, the inclusion of an ex-ante participation constraint for behavioral buyers (instead of interim participation) makes no difference since providing positive quality to the low valuation buyers is too costly, and the optimal mechanism here remains the same as the interim one: low valuations buyers are completely excluded and all high valuations buyers receive the same fully extracting contract.

We propose the following mechanism.

Proposition 3. *Fix θ_H, α and γ. Let $\overline{\theta} = \left(\frac{1}{2-\alpha}\right)\theta_H$ and $\underline{\theta} = \gamma\alpha\theta_H$.*

1. *For $\theta_L > \overline{\theta}$, the optimal mechanisms fully discriminates between behavioral and strategic buyers. For the strategic buyers, the standard mechanism without behavioral buyers is offered. For behavioral buyers the efficient quality is provided and their ex-ante expected rents are fully extracted.*

2. *For $\theta_L < \underline{\theta}$, the optimal mechanism offers the same contracts to both informational types, low valuation types are excluded, and high valuation types receive the efficient level of quality but have all their rents extracted.*

3. *For $\theta_L \in \left(\underline{\theta}, \overline{\theta}\right)$, there is a mechanism that satisfies ex-ante participation only for behavioral buyers which is incentive compatible and generates strictly more rents than the optimal interim mechanism.*

Proof. Consider the following mechanisms:

1. For $\theta_L > \frac{\theta_H}{2-\alpha}$,

$$q_L^S = \theta_L - \frac{\alpha}{1-\alpha}(\theta_H - \theta_L)$$
$$q_L^B = \theta_L$$
$$q_H^S = q_H^B = \theta_H$$
$$t_L^S = \theta_L q_L^S$$
$$t_L^B = [\alpha\theta_H + (1-\alpha)\theta_L]q_L^B$$
$$t_H^S = \theta_H^2 - (\theta_H - \theta_L)q_L^S$$
$$t_H^B \geq t_H^S$$
$$t_L^B \geq \theta_H\theta_L - (\theta_H - \theta_L)q_L^S$$
$$\alpha t_H^B + (1-\alpha)t_L^B = \alpha\theta_H^2 + (1-\alpha)\theta_L^2$$

2. For $\theta_L \in (\gamma\alpha\theta_H, \frac{\theta_H}{2-\alpha})$,

$$q_L^S = (1-\alpha)q_L^B$$
$$q_L^B = \frac{\theta_L - \gamma\alpha\theta_H}{\gamma(1-\alpha)^2 + (1-\gamma)}$$
$$q_H^S = q_H^B = \theta_H$$
$$t_L^S = \theta_L q_L^S$$
$$t_L^B = [\alpha\theta_H + (1-\alpha)\theta_L]q_L^B$$
$$t_H^S = t_H^B = \theta_H^2 - (\theta_H - \theta_L)q_L^S$$

3. For $\theta_L < \gamma\alpha\theta_H$,

$$q_L^S = q_L^B = t_L^S = t_L^B = 0$$
$$q_H^S = q_H^B = \theta_H$$
$$t_H^S = t_H^B = \theta_H^2$$

We will show that for each case the mechanisms above satisfy the required conditions in each case. *Case 1:* For high θ_L incentive compatibility and participation among strategic types follows from being the optimal mechanism in the standard screening problem without behavioral types. The last condition, $\alpha t_H^B + (1-\alpha)t_L^B = \alpha\theta_H^2 + (1-\alpha)\theta_L^2$, is equivalent to the ex-ante participation constraint for behavioral types.

Some algebra shows that incentive compatibility for type (H, S) with respect to (L, B) is equivalent to $t_L^B \geq \theta_H\theta_L - (\theta_H - \theta_L)q_L^S$. Moreover, from this inequality we can also conclude that $t_L^B \geq \theta_L^2$ since $q_L^S < \theta_L$. This implies incentive compatibility for (L, S) with respect to (L, B)

Finally, $q_H^S = q_H^B$ and $t_H^B \geq t_H^S$ implies incentive compatibility for strategic types with respect to (H, B).

It remains to show that this class of mechanisms generate higher profits than the interim mechanism. Indeed, we can show that they attain the maximum profits that could be generated in this context.

We do this by showing that the profits generated in this class of mechanisms from strategic types and behavioral types isolated are higher than the corresponding ones from the interim mechanism.

Notice that among strategic types, this class of mechanism indeed maximize the profits if we only consider incentive compatibility and participation among strategic types since they coincide with the optimal mechanism without behavioral types. Now, for behavioral, we know that this class of mechanism provides the efficient quality and extract all the utility from the behavioral types in expectation. Since we conclude that there is no way to increase revenue from neither group, these mechanism indeed achieve the maximum profit in this environment.

Case 2: For low θ_L, it coincides with the mechanism in Proposition 1, where all types receive zero utility, hence all participation constraints are binding. Moreover, any positive provision of quality to low valuation types reduces the profits that could be extracted from high types, and since $\theta_L < \gamma\alpha\theta_H$ this reduction effect dominates the revenue that could be obtained from low types.

Case 3: Let u_i^j be the equilibrium utility of a buyer of type (i, j). Then, $u_H^S = u_H^B = (\theta_H - \theta_L)q_L^S$, $u_L^S = 0$, and $u_L^B = -\alpha(\theta_H - \theta_L)q_L^B$.

From this is direct that no strategic type wants to deviate to the contract of his behavioral counterpart.

Moreover, the utility of a high valuation type under the contract of the low strategic type is equal to $(\theta_H - \theta_L)q_L^S$, while for imitating a low behavioral types it gets $(1-\alpha)(\theta_H - \theta_L)q_L^B = (\theta_H - \theta_L)q_L^S$. Hence, there is no profitable deviation for the high valuation type.

Finally, the utility of a low type under the contract of the high types is $-(\theta_H - \theta_L)(\theta_L - q_L^S)$, which is negative since $q_L^S < \theta_L$.

□

We conjecture that for $\theta_L \in (\underline{\theta}, \overline{\theta})$ the mechanism proposed in the proof of Proposition 3 is indeed optimal. While we were able to show by simulations that this in indeed the case, the formal proof is still pending.

We finish this section by stating a general property that any optimal mechanism should satisfy. As expected, all the proposed mechanisms in the proof of the proposition above satisfy this condition.

Claim 4. *In any optimal mechanism with ex-ante participation constraints only for the behavioral types, the participation constraint of behavioral types is binding.*

Proof. Suppose Γ is an optimal mechanism for which the participation constraint of the behavioral types is not binding. Consider an alternative mechanism $\hat{\Gamma}$ which coincide with Γ everywhere but in the transfers for behavioral types. In particular, $\hat{t}_L^B = t_L^B + \epsilon$ and $\hat{t}_H^B = t_H^B + \epsilon$, for $\epsilon > 0$ small enough such that the participation constraint is not violated.

Notice that this modified mechanism also satisfy the incentive compatibility constraints, doesn't violate any participation constraint, and generates strictly higher expected profits for the seller. Hence, Γ cannot be an optimal mechanism. □

6.2 Costly sophistication

Suppose now that our behavioral buyers could pay a fixed price f for becoming strategic *after learning their valuation types.* Hence, a behavioral buyer could pay to learn that he is behavioral. Here we are assuming that behavioral buyers are sophisticated and can perfectly calculate their expected payoff from all contracts but had self-control problems and cannot commit to not telling the truth after they have look into the contracts without some kind of help. In the context of the pricing example, we can think about this as some buyers looking for advice or doing some research in advance to figure out their optimal buying strategy. There, the difference between strategic and behavioral buyers will be that the first type already know the "tricks" used by the seller beforehand (for example, due to repeated interaction) while a behavioral buyer is exposed to those tricks unless he directly invest on preparing or informing himself.

Formally, we allow a behavioral buyer to become strategic paying a fee $f > 0$. We will refer to the act of becoming strategic as learning, and we will call a mechanism in which at least one behavioral type learns to be strategic as a *mechanism with learning.* Similarly, a mechanism where no buyer learns will be called a *mechanism without learning.*

The first step is to show that any mechanism with learning is weakly dominated by a mechanism without learning in terms of profits. That is, for any feasible mechanism with learning, there is a feasible mechanism without learning which weakly dominates it.

Lemma 2. *Let Γ be a feasible mechanism with learning. Then, there exist a feasible mechanism $\tilde{\Gamma}$ without learning such that the profits under $\tilde{\Gamma}$ are higher than under Γ.*

Proof. Consider a feasible mechanism Γ and suppose that type (i, B) learns in this mechanism. Consider a modified mechanism, $\tilde{\Gamma}$ such that $\Gamma_{-(q_i^B, t_i^B)} = \tilde{\Gamma}_{-(q_i^B, t_i^B)}$ and $(\tilde{q}_i^B, \tilde{t}_i^B) = (q_i^S, t_i^S + f)$. This modified mechanism is feasible and generates strictly more profits than the original mechanism. Moreover, under the modified mechanism, type (i, B) obtains the same payoff as in the original mechanism but without the need of learning. Hence, the new mechanism doesn't involve learning from (i, B).

Clearly, we could repeat the above procedure if there is another type (i', B) which learns under $\tilde{\Gamma}$ to construct a new mechanism without learning from this type. Thus, we can always construct a mechanism where no type learns. $\square$

Hence, it is without loss to restrict to mechanisms without learning. Here the assumption of commitment from the seller side is key since otherwise there is loss on considering mechanism without learning because the seller cannot commit to follow the mechanism without learning after the buyer choose not to learn.

The previous result lead us to define a set of *no learning constraints*,

$$\theta_i q_i^B - t_i^B \geq \theta_i q_i^S - t_i^S - f, \quad \forall i \in \{L, H\} \tag{NL}$$

Thus, we will look for an optimal mechanism in the class of feasible mechanisms that satisfy the no learning constraints.

Notice that the case where learning is not allowed is equivalent to a model with fee $f = \infty$.

The next proposition present the main result of this section.

Proposition 4. *Let* $\hat{\theta}_L = \left(\frac{\gamma\alpha}{\gamma\alpha+1-\alpha}\right)\theta_H$ *and* $\hat{f} = (\theta_H - \theta_L)\left(\theta_L - \gamma\left(\frac{\alpha}{1-\alpha}\right)(\theta_H - \theta_L)\right)$.

1. *If θ_L is either not bigger than $\hat{\theta}_L$, or bigger than $\hat{\theta}_L$ but the fee f is bigger than $\hat{f}$, then learning sophistication has no impact on the optimal contract.*

2. *Otherwise, if the fee is small enough learning could not be stopped and the optimal contracts imitates the one without behavioral types but where the behavioral high types must pay the fee to the seller instead of learning, or behavioral and strategic types face the same menu of contracts.*

3. *Also, if θ_L is big enough and the fee is neither too big nor too small, then a third case arises where the quality offered to low-behavioral types and low-strategic types is different, they are both lower than the efficient level and they are completely determined by the learning fee.*

Notice that the no learning constraint for low valuation types is never binding here since this types always get a payoff of zero in equilibrium regardless of their information status, so they have no incentives to learn.

Proof. We start by characterizing the optimal mechanism.

(i) If $\theta_L - \frac{\alpha}{1-\alpha}(\theta_H - \theta_L) > 0$ and

$$f > (\theta_H - \theta_L)\left(\theta_L - \frac{\gamma\alpha}{1-\alpha}(\theta_H - \theta_L)\right)$$

then

$$q_L^S = q_L^B = q_L = \theta_L - \frac{\gamma\alpha}{1-\alpha}(\theta_H - \theta_L)$$

$$q_H^S = q_H^B = \theta_H$$
$$t_H^S = \theta_H^2 - (\theta_H - \theta_L)q_L$$
$$t_H^B = \theta_H^2$$
$$t_L^S = t_L^B = \theta_L q_L$$

(ii) If $\theta_L - \frac{\alpha}{1-\alpha}(\theta_H - \theta_L) > 0$ and

$$f \in \left((\theta_H - \theta_L)\left(\theta_L - \frac{\alpha}{1-\alpha}(\theta_H - \theta_L)\right), (\theta_H - \theta_L)\left(\theta_L - \frac{\gamma\alpha}{1-\alpha}(\theta_H - \theta_L)\right)\right)$$

then

$$q_L^S = q_L^B = q_L = \frac{f}{\theta_H - \theta_L}$$
$$q_H^S = q_H^B = \theta_H$$
$$t_H^S = \theta_H^2 - (\theta_H - \theta_L)q_L$$
$$t_H^B = \theta_H^2$$
$$t_L^S = t_L^B = \theta_L q_L$$

(iii) If $\theta_L - \frac{\alpha}{1-\alpha}(\theta_H - \theta_L) > 0$ and

$$f < (\theta_H - \theta_L)\left(\theta_L - \frac{\alpha}{1-\alpha}(\theta_H - \theta_L)\right)$$

then

$$q_L^S = q_L^B = q_L = \theta_L - \frac{\alpha}{1-\alpha}(\theta_H - \theta_L)$$
$$q_H^S = q_H^B = \theta_H$$
$$t_H^S = \theta_H^2 - (\theta_H - \theta_L)q_L$$
$$t_H^B = \theta_H^2 - [(\theta_H - \theta_L)q_L - f]$$
$$t_L^S = t_L^B = \theta_L q_L$$

(iv) If $\theta_L \in \left(\frac{\gamma\alpha}{1-\alpha}(\theta_H - \theta_L), \frac{\alpha}{1-\alpha}(\theta_H - \theta_L)\right)$, and

$$f > (\theta_H - \theta_L)\left(\theta_L - \frac{\gamma\alpha}{1-\alpha}(\theta_H - \theta_L)\right)$$

then

$$q_L^S = q_L^B = q_L = \theta_L - \frac{\gamma\alpha}{1-\alpha}(\theta_H - \theta_L)$$

$$
\begin{aligned}
q_H^S = q_H^B &= \theta_H \\
t_H^S &= \theta_H^2 - (\theta_H - \theta_L)q_L \\
t_H^B &= \theta_H^2 \\
t_L^S = t_L^B &= \theta_L q_L
\end{aligned}
$$

(v) If $\theta_L \in \left(\frac{\gamma\alpha}{1-\alpha}(\theta_H - \theta_L), \frac{\alpha}{1-\alpha}(\theta_H - \theta_L)\right)$, and

$$f < (\theta_H - \theta_L)\left(\theta_L - \frac{\gamma\alpha}{1-\alpha}(\theta_H - \theta_L)\right)$$

then

$$
\begin{aligned}
q_L^S = q_L^B = t_L^S = t_L^B &= 0 \\
q_H^S = q_H^B &= \theta_H \\
t_H^S = t_H^B &= \theta_H^2
\end{aligned}
$$

(vi) If $\theta_L - \frac{\gamma\alpha}{1-\alpha}(\theta_H - \theta_L) < 0$

$$q_L^S = q_L^B = t_L^S = t_L^B = 0$$

$$
\begin{aligned}
q_H^S = q_H^B &= \theta_H \\
t_H^S = t_H^B &= \theta_H^2
\end{aligned}
$$

Now, we analyze each case of the proposition.

Case 1: This case is covered by points (i), (iv) and (vi). The proposed mechanism being optimal follows from the no learning constraints being trivially satisfied since the gains from learning are lower than the fee, and the proposed mechanism being the profit maximizing mechanism ignoring the no learning constraints.

Case 3: This case is covered by (ii). The mechanism in Case 1 violates the no learning constraints in this case. Hence, the no learning constraint must be binding. Clearly the proposed mechanism is feasible and satisfies the no learning conditions. Optimality follows from two observations. First, the profits generated under this mechanism are strictly higher than the profits generated by the mechanism which treats all types as strategic exactly when $\theta_L - \frac{\alpha}{1-\alpha}(\theta_H - \theta_L) > 0$ and $f < (\theta_H - \theta_L)\left(\theta_L - \frac{\alpha}{1-\alpha}(\theta_H - \theta_L)\right)$. Second, this mechanism is profit maximizing over the class of mechanism that assume the no learning condition for the high valuation type being binding.

Case 2: This case is covered by points (iii) and (v). Here, as in Case 3, the mechanism of Case 1 violates the learning condition but now we have also that the mechanism of Case 3 generates less profits than the mechanism which treats all types as strategic since the conditions for such dominance are violated. Since the proposed mechanism in this case is the mechanism which treats all agents as strategic, it is indeed optimal.

□

6.3 Constrained messages

In the previous sections we have assumed that any mechanism is available to the designer. In this section we will return to the basic model but assume instead that the designer faces some constraint in the "complexity" of contracts that could offer. In particular, we will restrict the set of available messages for the buyers. As we have shown above, the optimal mechanism for the unconstrained case involves using a menu with three different contracts. Hence, if the message space available for the designer contains at least three messages then the outcome of the optimal mechanism remains unchanged.

However, this is no longer the case if the designer can use less than three messages.

We can think about two different scenarios in this restricted environment: (i) the seller can only offer two different contracts, i.e., only two messages are available, or (ii) the designer can offer 3 contracts but one of those is fixed an equal to $(0, 0)$.[8]

For simplicity, we restrict to the first approach. The second approach was adopted by Li and Peters (2020) in the context of an auction environment[9].

We consider the same environment presented in Section 3 (interim participation, no learning possibility), but restricting the mechanisms available to the seller in terms of how many different contracts could be offered.

In particular, let $\Gamma = (\Gamma_1, \Gamma_2) = ((q_1, t_1), (q_2, t_2))$ be a mechanism in this context. That is, the mechanism used could offer at most two different contracts to the buyers. We will assume that (θ_L, B) always chooses Γ_1 while (θ_H, B) always chooses Γ_2. This assumption correspond to the translation

[8] We can also distinguish between at which point the outside option could be claimed: at the beginning or after an initial allocation is proposed. For the later, it implies that there could be some type of randomization in the sense that a contract could be offered knowing that only a particular type will accept while the other will reject it.

[9] Note that the model in Li and Peters (2020) differs in other dimensions as well. First, they consider the designing problem without appealing to the revelation principle directly and do not consider restricted message spaces as the one presented in this section. Also, they consider the full set of equilibria that could arise in the game between the designer and the agents, instead of the partial implementation or equilibrium selection approach adopted in our model.

of our original behavioral assumption over behavioral buyers to the current setting.

Notice that the revelation principle could not be applied directly in this environment since the set of potential messages is restricted[10].

Let $u(\theta, \sigma)$ be the payoff of an agent of type θ if he chooses contract 1 with probability σ (and contract 2 with probability $1 - \sigma$). As before, we will assume the buyers have quasilinear preferences, in consequence a buyer's payoff will take the form

$$u(\theta, \sigma) = \sigma (\theta q_1 - t_1) + (1 - \sigma) (\theta q_2 - t_2)$$

Let σ_i be the strategy of a strategic buyer with valuation θ_i. Then, the profits of the seller for a mechanism Γ and a strategy profile (σ_L, σ_H) is given by

$$\begin{aligned} R(\Gamma, \sigma_L, \sigma_H) = & [(1-\gamma)(1-\alpha) + \gamma\alpha\sigma_H) + \gamma(1-\alpha)\sigma_L] (t_1 - c(q_1)) \\ & + [(1-\gamma)\alpha + \gamma\alpha(1-\sigma_H) + \gamma(1-\alpha)(1-\sigma_L)] (t_2 - c(q_2)) \end{aligned}$$

Definition 5. *We say that the strategy profile (σ_L, σ_H) is a best response to the mechanism Γ if*

$$u(\theta_i, \sigma_i) \geq u(\theta_i, \sigma_i), \quad \forall i, \sigma_i'$$

Notice that this equivalent to the notion of incentive compatibility, taking into account the use of a strategy σ_i instead of truthful reporting.

Definition 6. *We say that Γ satisfies* participation *if*

$$u(\theta_L, 1) \geq 0 \text{ and } u(\theta_H, 0) \geq 0$$

That is, a mechanism satisfy participation if behavioral types are willing to accept their contract instead of exert their outside option (which is normalized to zero). As before, such participation constraints along with the incentive compatibility embedded in our best response requirement automatically imply that the participation conditions will hold for strategic buyers as well.

We can write the seller's problem as

$$\max_{\Gamma, \sigma_L, \sigma_H} R(\Gamma, \sigma_L, \sigma_H)$$

[10] It is possible to apply a modified revelation principle by introducing some constraints to the mechanisms which could be implemented. However, this approach would be equivalent to use the type of indirect mechanisms used here and does not necessarily provide a more tractable approach to study this problem.

subject to: (σ_L, σ_H) being a best response to Γ, and Γ satisfying participation.

Hence, the seller is not only choosing the mechanism but also the "equilibrium" played by the buyer. This characterization follows from considering only partial implementation instead of full implementation[11].

We will focus on equilibriums in pure strategies of this game.

6.3.1 Equilibrium

We start describing the equilibrium if the buyer is always behavioral. In this case, the optimal contract for the seller coincides with the "first best" or full information mechanism, and due to the behavioral assumption of this section is given by

$$\begin{aligned} q_1 &= \theta_L \\ t_1 &= \theta_L^2 \\ q_2 &= \theta_H \\ t_2 &= \theta_H^2 \end{aligned} \tag{5}$$

We refer to this mechanism as the *mechanism which ignores incentives* in this section.

This mechanism is implementable because the different messages/contracts that could be offered suffices to implement such scheme.

Now consider the other extreme case when the buyer is always strategic. In this case, the problem faced by the seller is the same as in the standard problem without behavioral types, hence the optimal mechanism for the seller is just the standard scheme when there is no behavioral buyers. Following the assumption we imposed for behavioral buyers, we let (strategic) buyers with low valuation choose contract Γ_1 and (strategic) buyers high valuation choose contract Γ_2, however if behavioral buyers are absent the labels of the contracts are arbitrary and flipping the definition of the contracts also constitutes an optimal mechanism here, while it is uniquely defined if there are behavioral buyers. We will refer to the mechanism below as the *mechanism which ignores behavioral buyers.*

$$\begin{gathered} q_1 = \max\left\{\theta_L - \tfrac{\alpha}{1-\alpha}\left(\theta_H - \theta_L\right), 0\right\} \\ q_2 = \theta_H \\ t_1 = \theta_L q_1 \\ t_2 = \theta_H^2 - (\theta_H - \theta_L) q_1 \end{gathered} \tag{6}$$

[11] The focus on partial implementation (equilibrium selection by the seller) was present on all the previous sections as well.

Again, since the optimal mechanism without behavioral types requires at most two different contracts, such contract is implementable in this environment, and hence optimal for the seller.

For intermediate cases, i.e., $\gamma \in (0,1)$, it can be shown that there is a unique $\hat{\gamma} \in (0,1)$ such that the first mechanism is optimal if $\gamma < \hat{\gamma}$, and the second one if $\gamma > \hat{\gamma}$.

Proposition 5. *Fix θ_L, θ_H and α. Then, there exist a unique $\hat{\gamma} \in (0,1)$ such that:*

1. *For $\gamma < \hat{\gamma}$ the mechanism which ignores incentives (5) is optimal.*
2. *For $\gamma > \hat{\gamma}$ the mechanism which ignores behavioral buyers (6) is optimal.*
3. *For $\gamma = \hat{\gamma}$ the mechanism which ignores incentives and the mechanism which ignores behavioral buyers generate the same profits for the seller, and both mechanisms are optimal.*

Proof. We start by characterizing the best response of the buyer.

We can restrict, without loss of generality, to contracts that satisfy $t_2 \geq t_1$ and $q_2 \geq q_1$. Then, best response for the buyer is as follows:

1. If $\theta_L(q_2 - q_1) > t_2 - t_1$, then $\sigma_L = \sigma_H = 0$
2. If $t_2 - t_1 > \theta_H(q_2 - q_1)$, then $\sigma_L = \sigma_H = 1$
3. If $\theta_H(q_2 - q_1) > t_2 - t_1 > \theta_L(q_2 - q_1)$, then $\sigma_L = 1$, $\sigma_H = 0$
4. If $\theta_H(q_2 - q_1) > t_2 - t_1 = \theta_L(q_2 - q_1)$, then $\sigma_L \in [0,1]$, $\sigma_H = 0$
5. If $\theta_H(q_2 - q_1) = t_2 - t_1 > \theta_L(q_2 - q_1)$, then $\sigma_L = 1$, $\sigma_H \in [0,1]$
6. If $\theta_H(q_2 - q_1) = t_2 - t_1 = \theta_L(q_2 - q_1)$, then $\sigma_L, \sigma_H \in [0,1]$

Furthermore, it is easy to verify that either 5 or 6 dominates in terms of revenue any other candidate mechanism. Also, we have that for γ close enough to zero, the contract which ignores incentive compatibility dominates, while for γ close enough to one 6 dominates. Finally, since profits under each contract are continuous and decreasing on γ, there is a unique point $\hat{\gamma}$ such that the mechanism in (5) dominates to the left and the mechanism in (6) dominates to the right. This proves the required result. □

We conclude this section presenting some comparative statics for this model.

Corollary 3. *Seller's profits are continuous and decreasing in γ. Welfare is decreasing in γ and discontinuous at $\hat{\gamma}$.*

Hence, both the seller and the society are benefited by the presence of behavioral buyers: since the seller could capture more rents from such buyers, this will induce him to provide a higher quality on average, which increases the social welfare. The discontinuity of the social welfare comes from the fact that the form of the optimal contract completely changes at $\hat{\gamma}$, while this doesn't have an impact on profits since the designer is optimally choosing to change the mechanism form, it has an impact on buyer's welfare. In particular, the quality received by a strategic buyer with high valuation increases from θ_L to θ_H, while the quality received by the buyers with low valuation types (both behavioral and strategic) is falls from θ_L to $\min\{0, \theta_L - \frac{\alpha}{1-\alpha}(\theta_H - \theta_L)\}$.

7 Concluding remarks

We have explored a monopolistic screening problem where the type of the buyer has two dimensions, his valuation and his informational type, and each one of them could have two values: either low or high valuation, and either strategic or behavioral. We have assumed that a behavioral buyer could not misreport his information, while a strategic buyer could misreport his information if he has incentives to doing so.

We have shown that the optimal mechanism in this context doesn't change too much with respect to the optimal mechanism without behavioral types. While, the exact form depends on the value of γ, the proportion of strategic types, as γ approaches 1, i.e., the standard model without behavioral types, the mechanism converges to the standard screening mechanism. This general characteristic holds even in the alternative environments considered in Section 6.

Further research should explore whether this results holds on more general settings, either by extending beyond profit maximization and the screening problem, but also on more rich behavior from behavioral types. In particular, whether those results hold if behavioral follows a different strategy instead of always report their true types truthfully.

References

Bergemann, Dirk and Stephen Morris (2005) "Robust Mechanism Design," *Econometrica*, 73 (6), 1771–1813, 10.1111/j.1468-0262.2005.00638.x.

[4]

Bobkova, Nina (2019) "Information Choice in Auctions." [5]

Börgers, Tilman (2015) *An introduction to the theory of mechanism design*, New York, NY: Oxford University Press. [3]

Carroll, Gabriel (2019) "Robustness in Mechanism Design and Contracting," *Annual Review of Economics*, 11 (1), 139–166, 10.1146/annurev-economics-080218-025616. [4]

de Clippel, Geoffroy, Rene Saran, and Roberto Serrano (2018) "Level-k Mechanism Design," *The Review of Economic Studies*, 86 (3), 1207–1227, 10.1093/restud/rdy031. [4]

Eliaz, Kfir (2002) "Fault Tolerant Implementation," *The Review of Economic Studies*, 69 (3), 589–610, 10.1111/1467-937X.t01-1-00023. [4]

Gleyze, Simon and Agathe Pernoud (2020) "Informationally Simple Incentives." [5]

Jehiel, Philippe, Benny Moldovanu, and Ennio Stacchetti (1999) "Multidimensional Mechanism Design for Auctions with Externalities," *Journal of Economic Theory*, 85 (2), 258 – 293, https://doi.org/10.1006/jeth.1998.2501. [3]

Li, Hao and Michael Peters (2020) "Unobserved Mechanisms." [3], [26]

Li, Shengwu (2017) "Obviously Strategy-Proof Mechanisms," *American Economic Review*, 107 (11), 3257–87, 10.1257/aer.20160425. [4]

Manelli, Alejandro M. and Daniel R. Vincent (2007) "Multidimensional mechanism design: Revenue maximization and the multiple-good monopoly," *Journal of Economic Theory*, 137 (1), 153 – 185, https://doi.org/10.1016/j.jet.2006.12.007. [3]

Ostrizek, Franz and Denis Shishkin (2019) "Screening with Frames." [4]

Rees-Jones, Alex and Samuel Skowronek (2018) "An experimental investigation of preference misrepresentation in the residency match," *Proceedings of the National Academy of Sciences*, 115 (45), 11471–11476, 10.1073/pnas.1803212115. [5]

ROCHET, JEAN-CHARLES (1987) "A necessary and sufficient condition for rationalizability in a quasi-linear context," *Journal of Mathematical Economics*, 16 (2), 191 – 200, https://doi.org/10.1016/0304-4068(87)90007-3. [3]

SALANT, YUVAL AND RON SIEGEL (2018) "Contracts with Framing," *American Economic Journal: Microeconomics*, 10 (3), 315–46, 10.1257/mic.20160230. [4]

SARAN, RENE (2011) "Bilateral trading with naive traders," *Games and Economic Behavior*, 72 (2), 544 – 557, https://doi.org/10.1016/j.geb.2010.09.009. [4], [8]

SEVERINOV, SERGEI AND RAYMOND DENECKERE (2006) "Screening when some agents are nonstrategic: does a monopoly need to exclude?" *The RAND Journal of Economics*, 37 (4), 816–840, 10.1111/j.1756-2171.2006.tb00059.x. [4]