PONTIFICIA
UNIVERSIDAD
CATÓLICA
DE CHILE

FACULTAD DE ECONOMÍA Y ADMINISTRACIÓN
INSTITUTO DE ECONOMÍA

Possibilities with mechanisms under limited commitment

by

Francisco Fuentes‡

Advisors : Nicolás Figueroa† and Tibor Heumann§

December, 2023

Santiago, Chile.

† PUC nicolasf@uc.cl
‡ PUC francisco.fuentes@uc.cl
§ PUC tibor.heumann@uc.cl

Abstract: This research is Francisco's MA thesis, the final report to qualify for a Master's degree in Economics. The main purpose of this document is to determine if a firm can capture value through information trading. We limit our efforts to a situation where a firm interacts with a customer in two stages but without any commitment, clearly, the agent is afraid to deliver private information because the firm can use it against him in future interactions and leave no informational rents. A relevant part of the analysis is the idea that a mechanism should encode not only the rules that determine the allocation but also the information available in each stage. We show that under a set of reasonable assumptions and different scenarios, the firm is actually incapable of selling information systems and it is limited to disclosing information through physical allocations. The constant trade-off between giving the agent the optimal menu and how much information the market is able to induce from it tomorrow seems to be enough trouble for the designer, so the absence of additional signals is the best option.

Keywords— Mechanism design, Limited commitment, Information design, Impossibility, Information markets

Contents

1. Introduction

1.1. Motivation

The information market has existed since time immemorial, with those adept at acquiring or safeguarding information often gaining a distinct advantage in future interactions. A prime illustration of such scenarios is evident in bargaining processes, where two agents seek to trade but are hesitant to reveal the true extent of their interest or the importance of the exchange. In any strategic game, the most effective player is one who can maintain a "poker face"; the absence of emotions can prove more advantageous than any physical, technological, or temporal privilege, consistently confounding rivals. This is not because any opponent is afraid of a confident face, but because they are not able to read any signal from other players forcing them to make the most predictable decision.

In a marketplace where goods and services are exchanged, physical allocations serve a dual purpose, also functioning as sources of information. The manner of dress, the choice of vehicle, or the possession of a home all convey signals regarding an individual's willingness to pay. Firms can strategically utilize this information in subsequent interactions. For example, Amazon can offer an even more personalized list of products the more you buy from their website. Netflix can predict what series you like based on how many hours you dedicate to different genres you have already seen. Any grocery online store such as Walmart is able to save your more frequent purchases to fill your cart for you periodically. Social media platforms like Instagram or TikTok consistently update their features, enticing users to disclose more information as they spend time on the platform, employing trial-and-error mechanics, every video you watch becomes a stage, and the more attention it receives marks the difference between success or failure, determining the next video, over and over again. Airlines offer different products based on past travels. Twitter, now called X, is also able to predict your interests based on your tweets and the tweets you have seen from others. Even software like operative systems, video games, virtual assistants, and web browsers adapt to your behavior and activities. Anyway, the list is endless, but they all share one common denominator, they make their predictions based on physical sources of information, measurable activities or observable purchases, any tangible interaction.

When considering pure information intermediation, illustrative examples are notably scarce. Twitch, an American video live-streaming service, operates a paid subscription system that, in essence, equates to little more than a logo alongside your username, indicating your level of fandom. In this scenario, payment is primarily driven by the desire to support the streamer and contribute to the growth of their community—akin to a donation. While subscribers may enjoy some preferential treatment, it's not guaranteed. In contrast, large-scale events like concerts, conventions, or festivals provide more tangible signals of preferential treatment, including guaranteed discounts and access to comfortable zones. Importantly, these interactions are rooted in commitment; when you purchase such benefits, assurance follows that you will receive them on the event day.

We focus on the most relevant aspects we have already discussed in the previous paragraphs and provide ideas explaining these scenarios with a classic dynamic mechanism design problem. The timing of the interactions, the information disclosed from them, and the limited commitment create an environment where pure information trading seems to be useless.

1.2. Example

This paper is looking for answers in a scenario where the designer is only able to commit in the short term, so the agent who participates is aware that any information revealed in the present stages can be used against them. The agent has fully persistent private information about their type, then naturally if they report it now the designer can remember it. Taking this into account, Doval and Skreta 2022 have studied the scenario, and to solve incentive compatibility issues they stated the mechanism designer has to consider the amount of information disclosed for future stages as part of the mechanism. So, input messages are always private and the designer is not able to see them, but we now consider output messages that represent how much information from past stages the principal uses in future stages. Then a classic mechanism which is a set of rules indicating which types receive which allocation now also involves how much information is disclosed through the allocation itself and signals.

The agent, having observed their private information ($\theta \in \Theta$), privately reports an input message, $m \in M$, to the mechanism, which then determines the distribution, $\varphi(.|m)$, from which an output message, $s \in S$, and an allocation, $a = (q, x) \in A$, are drawn. The output message and the allocation are publicly observable: they constitute the contractible parts of the mechanism.

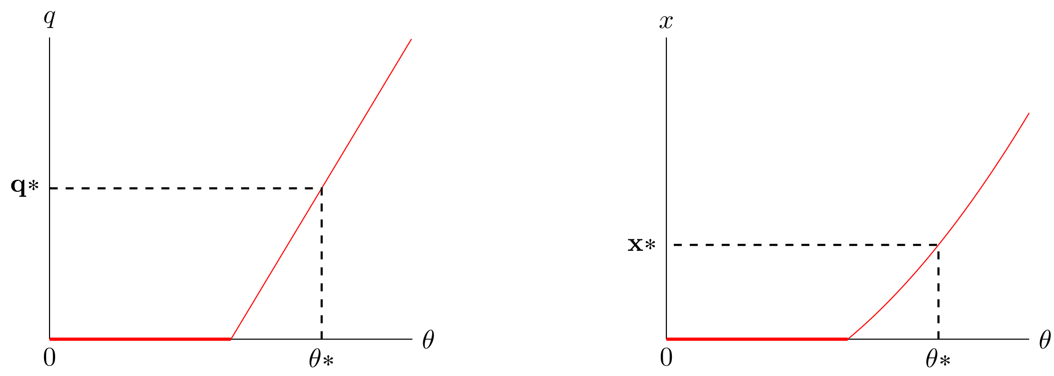

Fig. 1. Mussa and Rosen 1978 $\theta \sim U[0,1]$

Take figure 1 as an example of a possible mechanism in the first stage, classic Mussa and Rosen 1978 with a uniform distribution of types. If the designer observes $\mathbf{q}^*$ with $\mathbf{x}^*$, then it is clear their type is θ^*, basically the agent would make their private information completely public by accepting to participate in this mechanism. Therefore, an allocation rule that has invertible parts is also a full disclosure output message. However, those whose types are between 0 and 0.5 are indistinguishable from each other because they receive exactly the same allocation. We can say those who accept to participate and whose types are bigger than 0.5 certainly receive some allocation different from cero but they also public their information, while those whose types are less or equal to 0.5 may receive nothing in the first stage but at least they have some privacy. In this context, it is obvious that a mechanism like figure 1 is not incentive-compatible.

Note that signals $s \in S$ are completely irrelevant for types greater than 0.5 but they are certainly important for types less or equal to 0.5 since it is still possible to extract information from them. So, to include any information structure in our analysis and make it independent from the allocation is necessary to add signals that merely act as additional information.

Take a scenario where the allocation is actually constant, independent of the agent's type, this means signals could have big consequences and we could sell high types the possibility to be similar to low types sometimes. Even when this could be a good reason to add signals in the game, our research provides a formal proof in different scenarios where signals are actually worthless and they can not improve the designer's benefit in any form. There are several ideas behind the explanations, mostly based on the fact that any relevant signal system could be taken as a randomization of scenarios that could be replicated by the designer in future stages with no need for additional information. Moreover, the signal system must be consistent and it is impossible to induce convenient scenarios (low prices) for all types because the designer is aware there are different agents and low prices are not of equilibrium.

Formally we will show the principal only needs to use mechanisms of the form $\varphi : \Theta \to \Delta(A)$ because signals are not able to do any better. Allocations are enough sources of information and sometimes it is not necessary for the whole allocation to be observable, the result still holds. Agents are not willing to pay any additional for signal systems because mainly they decide how much information must be disclosed and agents do not want their private information to be public unless they receive a convenient allocation.

1.3. Literature review

A crucial aspect of our research involves mechanism design. We leverage examples from static mechanism design to underscore certain points, emphasizing they are not necessarily an equilibrium, particularly within our defined assumptions. The primary illustration stems from a classic corollary of Mussa and Rosen 1978, wherein a sole agent interacts with a single principal. Given our research's specific focus on such contexts, we deliberately exclude auctions from consideration since, by their nature, they involve multiple agents, akin to the scenario outlined by Myerson 1981. While the classic revelation principle is readily applicable to these scenarios, over subsequent decades, the community recognized the imperative for a more robust tool to address dynamic mechanism design. This arises not solely due to timing considerations but also due to the potential lack of commitment.

The natural extension in this literature lies in dynamic mechanism design, where several intriguing concepts arise exclusively from repeated interactions between an agent and a principal. Notably, the revelation principle becomes applicable due to the presence of full commitment, emerging as a valuable tool even in constrained scenarios. Myerson 1982 is particularly concerned with optimal coordination in mechanism design, directly employing incentive compatibility constraints. This same full commitment scenario is also demonstrated in the context of auction design by McAfee and Vincent 1997.

The significance of dynamic mechanism design research has grown over time, particularly due to the escalating need to relax the assumption of full commitment. The ability to negotiate in future stages becomes a pivotal element for designers, enabling the creation of contracts over time. However, certain contexts exist where a principal cannot commit to future stages, a situation acknowledged by both the designer and the agent, prompting them to anticipate the best outcomes for themselves in subsequent stages independent of past agreements. Moreover, information assumes an increasingly critical role; each interaction leaves valuable data for future stages. Consequently, the bargaining power is not necessarily constant over time, leading players to make more cautious decisions. The ratchet effect, highlighted in Freixas, Guesnerie, and Tirole 1985, becomes incredibly relevant, as irreversible actions carry significant consequences for the mechanism's output. In the absence of a revelation principle tailored for such scenarios, many papers adopt assumptions similar to Skreta 2006 or fix the mechanism, concentrating exclusively on game theory, as exemplified by Fudenberg and Tirole 1991. While various contributions exist, they often operate under scenarios where the problem is tractable enough to yield results.

In the near future, the first attempt to adapt the revelation principle to scenarios where it doesn't apply directly emerged. Bester and Strausz 2001 extends the principle to environments in which the mechanism designer cannot fully commit to the outcome induced by the mechanism. They show that he may optimally use a direct mechanism under which truthful revelation is an optimal strategy for the agent. In contrast with the conventional revelation principle, however, the agent may not use this strategy with probability one. Their results apply to contracting problems between a principal and a single agent. By reducing such problems to well-defined programming problems they provide a basic tool for studying imperfect commitment.

Certain texts in this context already incorporate an informational component, either directly integrated into the model or indirectly considered. When addressing private information, it is commonly treated as a primitive, given that each agent possesses private information about their types, which can be interpreted as an advantage, a value, or a preference. However, the extent to which this information becomes public in subsequent interactions generally arises as a consequence of the agent's actions and the principal's offers. While, in most cases, it is indeed a consequence, exploring it as a direct decision rather than a byproduct of other decisions proves to be an interesting avenue. This is where information design assumes significance as an approach. In the realm of pure information interactions, Kamenica and Gentzkow 2011 stands out as an ideal starting point. Furthermore, any subsequent contribution related to dynamic mechanism design cites Bayesian persuasion, acknowledging the natural informational aspect of the problem studied before, albeit without the detailed attention it merits. Numerous papers build on the ideas of Bayesian persuasion to address even more intricate scenarios, as exemplified by Bergemann, Brooks, and Morris 2015; Dworczak and Kolotilin 2019; Dworczak and Kolotilin 2019; Dworczak and Pavan 2022; Bergemann, Bonatti, and Gan 2022.

In the exploration of information design, several papers incorporate costs into the analysis, introducing robustness to the conclusions and extending the information problem to a more general scenario. Notably, Matyskova and Montes 2023 stands out as one of the most cited papers embracing this perspective, where the receiver can autonomously learn at a uniformly posterior-separable cost. Another noteworthy characterization is presented in Nguyen and Tan 2021, where the sender commits to a signal structure, privately observes the signal realization, and then sends a message to the receiver at a cost dependent on both the realized signal and the message sent. Across these models, a consistent and intuitive result emerges: the information designer (sender) is unable to achieve the optimal outcome described in Kamenica and Gentzkow 2011. While this outcome holds significant value, we think it will not substantially alter our research conclusions. In a strategic decision to maintain simplicity and avoid additional complexities in the analysis, we have opted to exclude costly information intermediation.

The most useful tool that we also apply in this research is the principle proved in Doval and Skreta 2022. They develop a tool akin to the revelation principle for dynamic mechanism-selection games in which the designer can only commit to short-term mechanisms. They identify a canonical class of mechanisms rich enough to replicate the outcomes of any equilibrium in a mechanism-selection game between an uninformed designer and a privately informed agent. A cornerstone of their methodology is the idea that a mechanism should encode not only the rules that determine the allocation but also the information the designer obtains from the interaction with the agent. Therefore, how much the designer learns, which is the key tension in design with limited commitment, becomes an explicit part of the design. Their result simplifies the search for the designer-optimal outcome by reducing the agent's behavior to a series of participation, truth-telling, and Bayes' plausibility constraints the mechanisms must satisfy.

In the context of industrial organizations, Doval and Skreta 2021 delves into a scenario where an agent interacts with either two firms or possibly a single firm across two stages. The particular interpretation is not of utmost importance; what truly matters is their utilization of a novel mechanism known as "Blackwell mechanism" to derive an optimal strategy within this context developed in Doval and Skreta 2022. This, together with ideas taken from Skreta 2006, are the most useful examples to argue why information must be taken as a direct decision variable to complete the analysis of limited commitment. Notably, the agent is concerned not only about the immediate interaction with a firm but also about the extent of information that will be disclosed to the firm in the subsequent stage as a result of this prior interaction. Some of their findings elucidate various pooling strategies, in which, at times, the optimal allocation as outlined in Myerson 1981 is disregarded. This is because, in the first stage, both the agent and the firm are willing to forego certain value in order to create market confusion in the subsequent stage.

In a far more general context Dworczak 2016 and Dworczak 2020 explain how we can take the incentive compatibility, individual rationality, and Bayes plausibility constraints and optimize a general objective function. However, he makes strong assumptions about the kind of mechanisms used, only limiting their results to "cutoff" allocation rules. In principle, this is not a big assumption because it is natural to think in an allocation rule where an agent gets the good if and only if their type is big enough. The result is in line with our findings because information disclosure seems to be useless in the presence of one agent. However, Dworczak 2016 has two different firms that interact with the agent, and Dworczak 2020 is more general, calling future interactions an aftermarket and not giving further details about it. This research is something in between those two scenarios, instead of two different firms we have only one firm with limited commitment, and instead of an aftermarket, the agent faces the same firm in future stages.

1.4. Organization

The rest of the paper is organized as follows. Section 2 describes the model and notation and Section 3 introduces the base theorem of impossibility, based on Dworczak 2016 ideas in a limited commitment scenario. Section 4 extends our analysis to a classic market instead of an information-only market. Section 5 discusses natural extensions and modifications to the problem. Finally, section 6 concludes.

2. The model

The narrative in this model unfolds as follows: a buyer (the agent) expresses a desire to purchase a specific quantity of a good and engages with a seller (the principal) over two consecutive days. Presently, the buyer possesses a type indicative of their valuation for the product offered by the seller. The seller, in turn, formulates an offer encompassing a defined quantity of the good, a charge (potentially positive or negative), and a signal system contingent on the buyer's type. While the quantity and charge are self-explanatory and may assume a probabilistic, lottery-like nature, the signal system might seem less intuitive initially. As clarified earlier, this signal system operates as a lottery that holds no direct payoff significance today but is correlated with the type revealed by the seller in the initial stage, thereby creating a spectrum of potential scenarios for tomorrow. Subsequently, this process iterates, taking into account all the information revealed the previous day. The allocation (comprising quantity and payments, punctually and functionally), the signal system, and the realized signal are all observable.

The model is an amalgam of the notation found in Doval and Skreta 2022 and ideas from Dworczak 2016. We need to compare and take ideas from both papers, this is the main objective behind this decision but even when the notation is the same as Doval and Skreta 2022 we have some differences.

The scenario we study is a subset of the scenarios studied in Doval and Skreta 2022 because we take into account only two stages instead of multiple stages or potentially infinite. This is mainly because we are trying to expand Dworczak 2016 in limited commitment scenarios. Moreover, this makes the analysts of the informational part easier to deal with allowing us to use backward induction directly.

We take into account only one player and one principal, this is less than Doval and Skreta 2022 and it is also less than Dworczak 2016 scenario but it is the same scenario of Doval and Skreta 2021. Note that our context is similar to Dworczak 2016, but it is not the same, Dworczak takes one agent who interacts with two different firms, and then incomes of the first stage do not affect the firm future incomes because they are completely different entities. In our research, the same firm interacts with the agent repeatedly, we did this intentionally to be closer to Doval and Skreta 2021 where the interaction is the same but the assumptions about types and agents are completely different and more specific.

2.1. Primitives

There's a principal (he) and the agent (she) who interact over two periods. The agent observes her type $\theta \in \Theta$ (discrete) distributed according to μ_1. Each period as a result of the interaction an allocation $a \in A$ and an output message $s \in S$ are determined. There are only two relevant allocations and we need to be specific that $a_t = (q_t, x_t) \in \mathbb{R}_+ \times \mathbb{R} \equiv A$ for $t \in \{1, 2\}$.

When the agent type is θ and the allocations are a_1 and a_2 then:

- The principal's payoff is $W(a_1, a_2, \theta)$
- The agent's payoff is $U(a_1, a_2, \theta)$

2.2. Mechanisms

In each period the allocation is determined by a mechanism $\mathbb{M}_t = (M_t, S_t, \varphi_t)$ with $t \in \{1, 2\}$.

- M_t input messages
- S_t output messages
- $\varphi_t : M_t \to \Delta(S_t \times A)$ assigns to each $m \in M_t$ a distribution over $S_t \times A$.

We can focus on mechanisms where $M_t = \Theta$ because of the revelation principle described in Doval and Skreta 2022.

2.3. Histories

Histories capture the mechanism ($\mathbb{M}_t$), the participation decision of the agent ($\pi_t \in \{0,1\}$), the output message (signal s_t), and the allocation (a_t). There are three relevant periods with histories and each one belongs to some space ($h^t \in H^t$ such that H^t is the cartesian product of all past possible combinations of interactions):

- $h^1 = \varnothing$
- $h^2 = (\mathbb{M}_1, \pi_1, s_1, a_1)$
- $h^3 = ((\mathbb{M}_1, \pi_1, s_1, a_1), (\mathbb{M}_2, \pi_2, s_2, a_2))$

There are some relevant observations to make here. First, histories are supposed to have all the game information and outcomes. Second, the history of the second period is key because it represents how information from the first period is disclosed through the mechanism $\mathbb{M}_1$, the signal or output message, and the allocation. Last but not least, π_t is supposed to be 1 in each period if a set of mechanisms is indeed implementable.

2.4. Beliefs

In the first stage of the game, the belief is the prior distribution μ_1. In the second stage, the belief μ_2 depends on h^2 just like Doval and Skreta 2022 describe. Then $\mu_2 : h^2 \to \Delta(\Theta)$ and the posterior specifically can be seen like:

$$\mu_2(\tilde{\theta}|h^2) = \mathbb{P}(\tilde{\theta}|\varphi_1, s_1, a_1) = \frac{\varphi_1(s_1, a_1|\tilde{\theta})\mu_1(\tilde{\theta})}{\sum_\theta \varphi_1(s_1, a_1|\theta)\mu_1(\theta)} \tag{B}$$

2.5. Strategies

Strategies could belong to a mechanism selection game.

- The principal: $\sigma_{Pt}(h^t) \in \Delta(\mathbb{M}_t)$ and $\sigma_P = (\sigma_{Pt})_{t=1}^2$
- The agent: $\sigma_{At}(\theta, h^t, \mathbb{M}_t) \in \Delta((\pi_t, r_t))$ where $\pi(\theta, h^t, \mathbb{M}_t)$ represent the agent's participation decision and $r_t(\theta, h^t, \mathbb{M}_t)$ describes the agent's choice of input messages in the mechanism given she participates.
- $(\sigma_P, \sigma_A, \mu)$ defines an assessment.

We do not need to focus on mechanism selection games because we are working under the assumption that $M_t = \Theta$ and the signal space is fixed S. Then the only thing we are selecting is the best allocation rule $\varphi_t : \Theta \to \Delta(S \times A)$ in the space of direct mechanisms which is without loss of generality. So, σ_{Pt} indicates a direct mechanism with probability 1, moreover, we can focus on mechanisms with participation $\pi_t = 1$ and r_t indicates the type with probability 1 (of course, each one represents a constraint in the optimization problem).

2.6. The problem

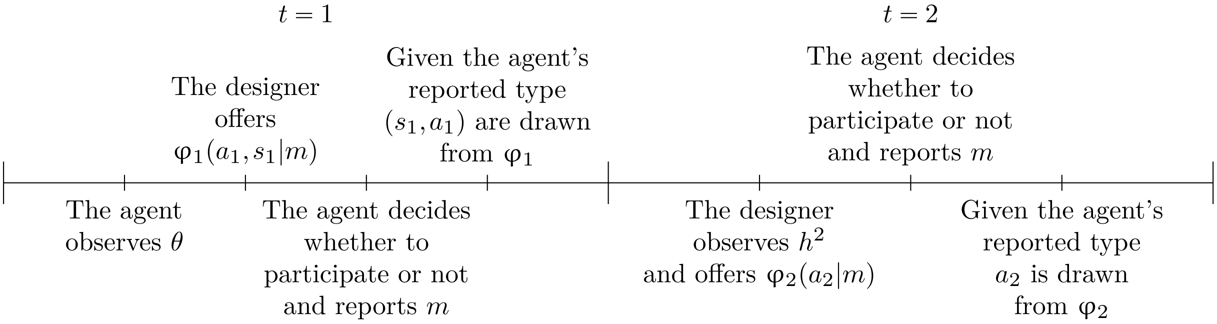

Fig. 2. Timing of the game

Like any other dynamic game, the problem can be solved by backward induction. **The second stage scenario** would be similar to Dworczak 2016 third-party problem, given some h^2:

$$\max_{\varphi_2} \mathbb{E}_{\theta \sim \mu_2(h^2)} \left[\int_{a_2} W(a_1, a_2, \theta) \varphi_2(da_2|\theta) \right]$$

s.t.

$$\int_{a_2} U(a_1, a_2, \theta) \varphi_2(da_2|\theta) \geqslant 0 \qquad \forall \theta \in \Theta \tag{II-IR}$$

$$\int_{a_2} U(a_1, a_2, \theta) \varphi_2(da_2|\theta) \geqslant \int_{a_2} U(a_1, a_2, \theta) \varphi_2(da_2|\tilde{\theta}) \qquad \forall \tilde{\theta} \neq \theta \tag{II-IC}$$

Let $\varphi_2 : H^2 \times \Theta \to \Delta(A)$ be a solution to the problem above. We are going to write any solution to the problem as $\varphi_2^{s_1, a_1}$ where (s_1, a_1) are the observed realization of the signal and allocation of the first stage respectively. We also know that $h^2 \in H^2$ actually has a lot more information and the solution $\varphi_2^{s_1, a_1}$ depends punctually and functionally on the first stage decisions, so we are abusing notation.

From now on, we will call $\mathcal{U}^{s_1, a_1}(\theta)$ to the informational rents obtained by the agent of type θ who participates in $\varphi_2^{s_1, a_1}$ given some h^2. Then, the **first stage problem** or the equivalent to the mediator problem described in Dworczak 2016 is constrained by:

$$\int_{s_1} \int_{a_1} \mathcal{U}^{s_1, a_1}(\theta) \varphi_1(da_1, ds_1|\theta) \geqslant 0 \qquad \forall \theta \in \Theta \tag{I-IR}$$

$$\int_{s_1} \int_{a_1} \mathcal{U}^{s_1, a_1}(\theta) \varphi_1(da_1, ds_1|\theta) \geqslant \int_{s_1} \int_{a_1} \mathcal{U}^{s_1, a_1}(\theta) \varphi_1(da_1, ds_1|\tilde{\theta}) \qquad \forall \tilde{\theta} \neq \theta \tag{I-IR}$$

The main objective of the next section of the research is to study which mechanisms are implementable, therefore we do not need to explicitly model the first-stage interests even when those are given by $W(.)$, in other words, the principal's objective in the first stage is irrelevant. In section 4, the principal's objective function will be particularly relevant.

3. The relevance of an allocation which is payoff irrelevant

More specifically, we are not going to take the whole allocation as payoff irrelevant but we will check how much the firm can do when the quantity offered is not important for the agent. This, in principle, does not mean the principal has run out of options to trade in the first stage because we have to remember he still can sell signal systems.

So, as a first approach, we are going to study a scenario of information trading in the first stage. There is a principal (firm) who interacts with an agent (customer) over two periods. The agent does not care about the good in the first stage but she can be charged anyway for a potentially benefit signal system which induces a variety of mechanisms in the second stage. We will show that any signal system is actually worthless because the principal lacks the ability to influence himself effectively.

Definition: Given some mechanism $\varphi_1 : \Theta \to \Delta(S \times A)$ together with $\varphi_2 : H^2 \times \Theta \to \Delta(A)$ we will call an outcome:

1. The distribution over q_2 given some θ.

$$\int_{s_1} \int_{a_1} \int_{x_2} \varphi_2^{s_1,a_1}(a_2|\theta)\varphi_1(da_1, ds_1|\theta)dx_2$$

2. The distribution over q_1 given some θ.

$$\int_{s_1} \int_{x_1} \varphi_1(a_1, ds_1|\theta)dx_1$$

3. The expected payment given some θ.

$$\int_{s_1} \int_{a_1} \left[x_1 + \int_{a_2} x_2 \varphi_2^{s_1,a_1}(da_2|\theta) \right] \varphi_1(da_1, ds_1|\theta)$$

The first result of our research can be seen as proof which aims to state that two sets are the same. The direct way to do it consists of realizing the set of outcomes that can be induced by the firm and the set of outcomes produced by no information disclosure are subsets of each other.

We can see the set of all possible outcomes as a set of tuples of the three elements described in the previous definition, we call this set $\mathbb{O} \subset \Delta(\mathbb{R}_+) \times \mathbb{R}$ (note we only care about the distribution over q_2 in this case). And we can take the set of outcomes induced by no information as $\mathbb{O}_\varnothing$. So it must be true that:

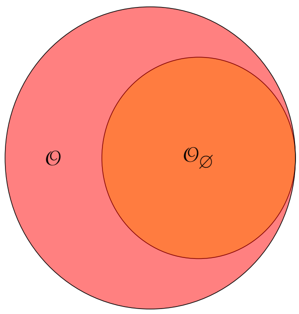

Fig. 3. No information is subset of information

Theorem 1 proves the unexpected part, that actually $\mathbb{O} \subset \mathbb{O}_\varnothing$ and therefore they are the same set. This is essentially the same process followed in Dworczak 2016.

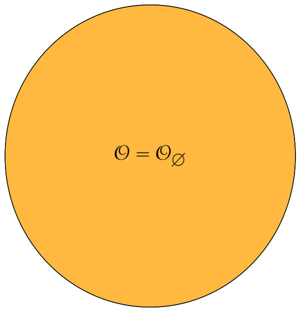

Fig. 4. No information is the same as information

3.1. *Additional assumptions*

1. Separability utility functions.

   The agent's utility function can be written with transfers that impact only linearly.

$$U(a_1, a_2, \theta) = u_1(q_1, \theta) - x_1 + u_2(q_2, \theta) - x_2 \tag{1.A.1}$$

   The principal's payoff function can be written as:

$$W(a_1, a_2, \theta) = x_1 - c_1(q_1) + x_2 - c_2(q_2) \tag{1.A.2}$$

2. The quantity allocated in period one (q_1) is payoff irrelevant.

   Specifically, $u_1(q_1, \theta) = 0$ and $c_1(q_1) = 0$. This is an exploratory assumption, a benchmark to start with. We will drop it in later sections.

$$U(a_1, a_2, \theta) = u_2(q_2, \theta) - x_2 - x_1 \tag{1.A.3}$$

$$W(a_1, a_2, \theta) = x_1 + x_2 - c_2(q_2) \tag{1.A.4}$$

3.2. *The problem*

Given the fact that a_1 is fixed in the second stage, we can write the agent utility as:

$$\begin{aligned}\int_{a_2} U(a_1, a_2, \theta)\varphi_2(da_2|\theta) &= \int_{a_2} [u_2(q_2, \theta) - x_2 - x_1]\varphi_2(da_2|\theta) \\ &= \int_{a_2} u_2(q_2, \theta)\varphi_2(da_2|\theta) - \bar{x}_2(\theta) - x_1\end{aligned}$$

Where:

$$\bar{x}_2(\theta) = \int_{a_2} x_2 \varphi_2(da_2|\theta)$$

Note that because the transfers affect the utility with a linear behavior the agent only cares about the average payment instead of the whole distribution.

Now, the **Second stage problem** can be rewritten.

$$\max_{\varphi_2} \mathbb{E}_{\theta \sim \mu_2(h_2)} \left[\int_{a_2} W(a_1, a_2, \theta)\varphi_2(da_2|\theta) \right]$$

s.t.

$$\int_{a_2} u_2(q_2,\theta)\varphi_2(da_2|\theta) - \bar{x}_2(\theta) \geqslant 0 \qquad \forall\theta \in \Theta \tag{1.II-IR}$$

$$\int_{a_2} u_2(q_2,\theta)\varphi_2(da_2|\theta) - \bar{x}_2(\theta) \geqslant \int_{a_2} u_2(q_2,\theta)\varphi_2(da_2|\tilde{\theta}) - \bar{x}_2(\tilde{\theta}) \qquad \forall\tilde{\theta} \neq \theta \tag{1.II-IC}$$

Consider now the **First stage problem** where the agent's future utility under truth-telling messages, given some $\varphi_2^{s_1,a_1} : \Theta \to \Delta(A)$ a solution to the problem above, is:

$$\mathcal{U}^{s_1,a_1}(\theta) = \int_{a_2} U(a_1,a_2,\theta)\varphi_2^{s_1,a_1}(da_2|\theta) = \int_{a_2} u_2(q_2,\theta)\varphi_2^{s_1,a_1}(da_2|\theta) - \bar{x}_2(\theta) - x_1$$

The first stage expected utility is:

$$\int_{s_1}\int_{a_1} \mathcal{U}^{s_1,a_1}(\theta)\varphi_1(da_1,ds_1|\theta) = \int_{s_1}\int_{a_1}\left[\int_{a_2} u_2(q_2,\theta)\varphi_2^{s_1,a_1}(da_2|\theta) - \bar{x}_2(\theta) - x_1\right]\varphi_1(da_1,ds_1|\theta)$$

Note that we can rewrite this whole problem by changing the integration order. For example, take only the utility function separately.

$$\int_{s_1}\int_{a_1}\left[\int_{a_2} u_2(q_2,\theta)\varphi_2^{s_1,a_1}(da_2|\theta)\right]\varphi_1(da_1,ds_1|\theta)$$

We know this is the same as:

$$\int_{a_2}\int_{s_1}\int_{a_1} u_2(q_2,\theta)\varphi_2^{s_1,a_1}(da_2|\theta)\varphi_1(da_1,ds_1|\theta) = \int_{a_2} u_2(q_2,\theta)\int_{s_1}\int_{a_1} \varphi_2^{s_1,a_1}(da_2|\theta)\varphi_1(da_1,ds_1|\theta)$$

Then we can define any allocation rule potentially implemented in the second stage induced by the first stage interaction as:

$$\phi(.|\theta) = \int_{s_1}\int_{a_1} \varphi_2^{s_1,a_1}(da_2|\theta)\varphi_1(da_1,ds_1|\theta)$$

This is exactly the same logic explained in Dworczak 2016. Again using the fact that the distribution of both transfers does not matter because the agent's only concern is the final average we can declare the next variables to make the notation simpler.

$$X_1(\theta) = \int_{s_1}\int_{a_1} x_1\varphi_1(da_1,ds_1|\theta)$$

$$X_2(\theta) = \int_{a_2}\int_{s_1}\int_{a_1} x_2\varphi_2^{s_1,a_1}(da_2|\theta)\varphi_1(da_1,ds_1|\theta) = \int_{s_1}\int_{a_1} \bar{x}_2(\theta)\varphi_1(da_1,ds_1|\theta) = \int_{a_2} x_2\phi(da_2|\theta)$$

Then a mechanism $\varphi_1 : \Theta \to \Delta(S \times A)$, which induces a mechanism $\phi(a_2|\theta)$ in the second stage, is implementable as long as:

$$\int_{a_2} u_2(q_2,\theta)\phi(da_2|\theta) - X_2(\theta) - X_1(\theta) \geqslant 0 \qquad \forall\theta \in \Theta \tag{1.I-IR}$$

$$\int_{a_2} u_2(q_2,\theta)\phi(da_2|\theta) - X_2(\theta) - X_1(\theta) \geqslant \int_{s_1}\int_{a_1} \mathcal{U}^{s_1,a_1}(\theta)\varphi_1(da_1,ds_1|\tilde{\theta}) \qquad \forall\tilde{\theta} \neq \theta \tag{1.I-IC}$$

Now the transfer is also part of the information disclosed to the second stage while in Dworczak 2016 the signal is the only source of information the third party can see and nothing more. This makes the transfer, the only relevant part of the allocation in the first stage, a variable that needs to be decided together with the signal system.

3.3. First Theorem

Theorem 1

The firm cannot influence the outcome of the second stage by releasing information of any form. This is, the principal is able to induce the same mechanisms in the second stage using mechanisms of the form $\varphi_1 : \Theta \to \Delta(S \times A)$ as he can do with no mechanisms at all.

Theorem 1 intuition is driven by the fact that "high types" are always looking for signals that implement low prices tomorrow, however, "low types" barely trade and if they do they also need prices to be low so they can be at least indifferent between participating or not. The principal in the second stage is able to predict this obvious outcome and therefore low prices are not an equilibrium strategy. If any combination of signals is implementable then it is not that informative about the customer type.

Proof

Consider $\varphi_1 : \Theta \to \Delta(S \times A)$ an implementable mechanism in the first stage, which induces $\phi : \Theta \to \Delta(A)$, $X_2 : \Theta \to \mathbb{R}$ and $X_1 : \Theta \to \mathbb{R}$, then it must be true that:

The whole mechanism must be incentive-compatible.

$$\int_{a_2} u_2(q_2,\theta)\phi(da_2|\theta) - X_2(\theta) - X_1(\theta) \geqslant \int_{a_2} u_2(q_2,\theta)\phi(da_2|\tilde{\theta}) - X_2(\tilde{\theta}) - X_1(\tilde{\theta}) \qquad \forall\tilde{\theta} \neq \theta \tag{1.P.1}$$

The agent must participate.

$$\int_{a_2} u_2(q_2,\theta)\phi(da_2|\theta) - X_2(\theta) - X_1(\theta) \geqslant 0 \tag{1.P.2}$$

The principal is better off with the information than without it in the second stage.

$$\mathbb{E}_\theta\left[\int_{s_1}\int_{a_1}\int_{a_2} W(a_1,a_2,\theta)\varphi_2^{s_1,a_1}(da_2|\theta)\varphi_1(da_1,ds_1|\theta)\right] \geqslant v^* \tag{1.P.3}$$

Defining $\hat{X}(\theta) = X_2(\theta) + X_1(\theta)$ then the expression above tells us that any distribution, such that its transfers mean is $\hat{X}(\theta)$ and its marginal distribution in q_2 is the same as ϕ, satisfies 1.II-IR in the second stage. For example an allocation rule $\eta : \Theta \to \Delta(A)$ such that:

$$\int_{x_2} \eta(a_2|\theta)dx_2 = \int_{x_2} \phi(a_2|\theta)dx_2 \qquad \forall\theta \in \Theta \tag{1.P.4}$$

$$\int_{a_2} x_2\eta(da_2|\theta) = \hat{X}(\theta) \qquad \forall\theta \in \Theta \tag{1.P.5}$$

Now we can see that η also satisfies 1.II-IC because of 1.P.1:

$$\int_{a_2} u_2(q_2,\theta)\phi(da_2|\theta) - \hat{X}(\theta) \geqslant \int_{a_2} u_2(q_2,\theta)\phi(da_2|\tilde{\theta}) - \hat{X}(\tilde{\theta}) \qquad \forall\tilde{\theta} \neq \theta$$

Then η is available even when there is no information disclosed in the first stage and:

$$\mathbb{E}_\theta\left[\int_{a_2} W(a_1,a_2,\theta)\eta(da_2|\theta)\right] \leqslant v^* \tag{1.P.6}$$

Using 1.P.5, 1.P.4 and 1.A.4 in 1.P.6 together with the fact that if no information was disclosed in the first stage then x_1 would be a constant, specifically $x_1 = 0$, then:

$$\mathbb{E}_\theta\left[\int_{a_2} W(a_1,a_2,\theta)\eta(da_2|\theta)\right] = \mathbb{E}_\theta\left[\int_{a_2} [x_2 - c_2(q_2)]\eta(da_2|\theta)\right] = \mathbb{E}_\theta\left[\hat{X}(\theta) - \int_{a_2} c_2(q_2)\eta(da_2|\theta)\right]$$
$$= \mathbb{E}_\theta\left[\hat{X}(\theta) - \int_{a_2} c_2(q_2)\phi(da_2|\theta)\right] \leqslant v^*$$

Moreover, using 1.A.4 in 1.P.3:

$$\mathbb{E}_\theta\left[\int_{s_1}\int_{a_1}\int_{a_2} W(a_1,a_2,\theta)\varphi_2^{s_1,a_1}(da_2|\theta)\varphi_1(da_1,ds_1|\theta)\right] = \mathbb{E}_\theta\left[\int_{s_1}\int_{a_1}\int_{a_2} [x_1 + x_2 - c_2(q_2)]\varphi_2^{s_1,a_1}(da_2|\theta)\varphi_1(da_1,ds_1|\theta)\right]$$
$$= \mathbb{E}_\theta\left[X_1(\theta) + X_2(\theta) - \int_{s_1}\int_{a_1}\int_{a_2} c_2(q_2)\varphi_2^{s_1,a_1}(da_2|\theta)\varphi_1(da_1,ds_1|\theta)\right]$$
$$= \mathbb{E}_\theta\left[X_1(\theta) + X_2(\theta) - \int_{a_2} c_2(q_2)\phi(da_2|\theta)\right] \geqslant v^*$$

We know that $\hat{X}(\theta) = X_1(\theta) + X_2(\theta)$, therefore:

$$v^* \leqslant \mathbb{E}_\theta\left[X_1(\theta) + X_2(\theta) - \int_{a_2} c_2(q_2)\phi(da_2|\theta)\right] = \mathbb{E}_\theta\left[\hat{X}(\theta) - \int_{a_2} c_2(q_2)\phi(da_2|\theta)\right] \leqslant v^*$$

Finally, we conclude that what the principal can do without information disclosure is the same as what the principal can do with it. ■

4. No additional information needed

A more intuitive scenario would be one where the principal offers both, a relevant physical allocation and a signal system. Now, the question is different, because the allocation by itself is really necessary to improve the mechanism's outcome and therefore is very unlikely that a context completely free of information could have the same possibilities as one somehow informed. Note this is not due to the information being relevant to improve the scenario, but because the allocation acts as an informative outcome and can not be avoided by the designer. Then, we now focus on answering if the firm is able to improve its benefits through information disclosure beyond the one disclosed by the allocation itself. Theorem 2 addresses this question directly and it shows how limited the principal is, even when he is able to use a signal, mostly because any form of details about the agent will be public tomorrow only as long as she willingly accepts it (as long as the signal is incentive-compatible).

In this scenario, the principal's tools are more than merely information but also act as signals because allocated quantities are also observable. However, the question changes now because we focus on answering how significant is pure information intermediation, when the principal updates his belief based on physical allocations it is not cataloged as pure information because these sources were actually payoff-relevant in the previous stage.

The primary finding in this section is also grounded in the concept of outcomes, yet it represents a less general result compared to Theorem 1. However, its utility should not be underestimated, as its specificity serves a distinct purpose. So, take $\mathbb{O}_{s(\varnothing)}$ as the set of all possible outcomes when the first stage mechanism does not use signals (or it uses non-informative signals) and $\mathbb{O}$ as all possible outcomes. Moreover, the optimal output for the principal is represented by o^*.

We know $o^* \in \mathbb{O}$ but we do not know if it is exclusive of this set or, perhaps, it belongs to $\mathbb{O}_{s(\varnothing)}$ as figure 5 show us.

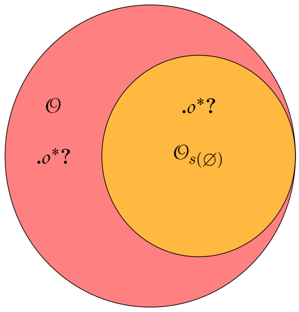

Fig. 5. The optimal outcome in the space

Theorem 2 concludes that o^* actually belongs in $\Theta_{s(\varnothing)}$ so it is without loss of generality to focus only on mechanisms with no signal systems when the objective is to maximize the principal's utility just like figure 6 represents.

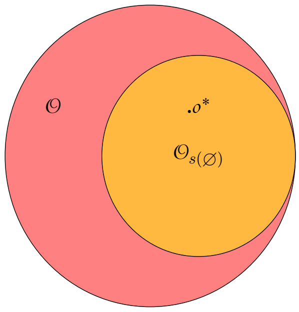

Fig. 6. The optimal outcome is in the smallest set

4.1. Additional assumptions

1. Separability utility functions.

   The agent's utility function can be written with transfers that impact only linearly.

$$U(a_1, a_2, \theta) = u_1(q_1, \theta) - x_1 + u_2(q_2, \theta) - x_2 \tag{2.A.1}$$

   The principal's payoff function can be written as:

$$W(a_1, a_2, \theta) = x_1 - c_1(q_1) + x_2 - c_2(q_2) \tag{2.A.2}$$

2. Deterministic mechanism in the first stage

   We will work under the assumption that the mechanism in the first stage has deterministic allocation rules. So, a mechanism in the first stage is $\varphi_1 : \Theta \to A \times \Delta(S)$ and it can be written as a tuple $(a_1(\theta), \gamma(./\theta))$ where $a_1 : \Theta \to A$ and $\gamma : \Theta \to \Delta(S)$. This is because the analysis becomes relatively simpler than a scenario where any randomization is available. The assumption is not without loss of generality but we hope to relax it in future iterations.

4.2. The problem

Taking into account mechanisms $\varphi_2 : \Theta \to \Delta(A)$, the **Second stage problem** given some h^2 is:

$$\max_{\varphi_2} \mathbb{E}_{\theta\sim\mu_2(h^2)}\left[\int_{a_2}(x_2 - c_2(q_2))\varphi_2(da_2|\theta)\right]$$

s.t.

$$\int_{a_2}(u_2(q_2,\theta) - x_2)\varphi_2(da_2|\theta) \geqslant 0 \qquad \forall\theta\in\Theta \tag{2.II-IR}$$

$$\int_{a_2}(u_2(q_2,\theta) - x_2)\varphi_2(da_2|\theta) \geqslant \int_{a_2}(u_2(q_2,\theta) - x_2)\varphi_2(da_2|\tilde{\theta}) \qquad \forall\tilde{\theta}\neq\theta \tag{2.II-IC}$$

Let $\varphi_2 : H^2 \times \Theta \to \Delta(A)$ be a solution to the problem above. The **First stage problem** is:

$$\max_{a_1,\gamma} \mathbb{E}_{\theta}\left[x_1(\theta) - c_1(q_1(\theta)) + \int_{s_1}\left(\int_{a_2}[x_2 - c_2(q_2)]\varphi_2^{s_1,a_1(\theta)}(da_2|\theta)\right)\gamma(ds_1|\theta)\right]$$

s.t.

$$u_1(q_1(\theta),\theta) - x_1(\theta) + \int_{s_1}\left(\int_{a_2}[u_2(q_2,\theta) - x_2]\varphi_2^{s_1,a_1(\theta)}(da_2|\theta)\right)\gamma(ds_1|\theta) \geqslant 0 \qquad \forall\theta\in\Theta \tag{2.I-IR}$$

$$\begin{aligned} u_1(q_1(\theta),\theta) - x_1(\theta) + \int_{s_1}\left(\int_{a_2}[u_2(q_2,\theta) - x_2]\varphi_2^{s_1,a_1(\theta)}(da_2|\theta)\right)\gamma(ds_1|\theta) \geqslant \\ \int_{a_1}\int_{s_1}\mathcal{U}^{s_1,a_1}(\theta)\varphi_1(ds_1,da_1|\tilde{\theta}) \qquad \forall\tilde{\theta}\neq\theta \end{aligned} \tag{2.I-IC}$$

4.3. Second Theorem

Theorem 2

The firm cannot improve its profits by releasing additional information through signals in the first stage. In other words, the principal's maximum utility when he uses mechanisms of the form $\varphi_1 : \Theta \to A \times \Delta(S)$ together with $\varphi_2 : H^2 \times \Theta \to \Delta(A)$ is the same as when he is limited to mechanisms of the form $\varphi_1 : \Theta \to A$ with $\varphi_2 : H^2 \times \Theta \to \Delta(A)$.

Proof

Consider $\varphi_1(\theta) = (a_1(\theta), \gamma(.|\theta))$ and its induced mechanisms in the second stage, as a **solution** to the problem. Then it must be true that:

The whole mechanism is incentive-compatible, in particular, a constant report through stages does not represent a profitable deviation.

$$\begin{aligned} u_1(q_1(\theta),\theta) - x_1(\theta) + \int_{s_1}\left(\int_{a_2}[u_2(q_2,\theta) - x_2]\varphi_2^{s_1,a_1(\theta)}(da_2|\theta)\right)\gamma(ds_1|\theta) \geqslant \\ u_1(q_1(\tilde{\theta}),\theta) - x_1(\tilde{\theta}) + \int_{s_1}\left(\int_{a_2}[u_2(q_2,\theta) - x_2]\varphi_2^{s_1,a_1(\tilde{\theta})}(da_2|\tilde{\theta})\right)\gamma(ds_1|\tilde{\theta}) \qquad \forall\tilde{\theta}\neq\theta \end{aligned} \tag{2.P.1}$$

The agent must willingly participate in both stages.

$$u_1(q_1(\theta),\theta)-x_1(\theta)+\int_{s_1}\left(\int_{a_2}[u_2(q_2,\theta)-x_2]\varphi_2^{s_1,a_1(\theta)}(da_2|\theta)\right)\gamma(ds_1|\theta)\geqslant 0 \qquad \forall\tilde{\theta}\neq\theta \tag{2.P.2}$$

$$\int_{a_2}[u_2(q_2,\theta)-x_2]\varphi_2^{s_1,a_1(\theta)}(da_2|\theta)\geqslant 0 \qquad \forall\tilde{\theta}\neq\theta \tag{2.P.3}$$

Consider the next mechanisms for both stages:

1. In the first stage we are going to use the same allocation $((a_1(\theta),-))$ as the original mechanism but without any signal $(\tilde{\varphi}_1(\theta)=a_1(\theta))$.
2. Given some history h^2 (an allocation a previously observed), and a message $\tilde{\theta}$ in the second stage, the proposed mechanism is:

$$\tilde{\varphi}_2^a(a_2|\tilde{\theta})=\begin{cases}\int_{s_1}\varphi_2^{s_1,a}(a_2|\tilde{\theta})\gamma(ds_1|\tilde{\theta}) & \text{if } a_1(\tilde{\theta})=a\\ \Omega(a_2) & \text{if } a_1(\tilde{\theta})\neq a\end{cases} \tag{2.P.4}$$

   Where $\Omega(a_2)=1$ when $a_2=(0,\infty)$ and $\Omega(a_2)=0$ otherwise.

Proposition 2.P.5: The mechanism $\tilde{\varphi}_2^a$, given some truthful h^2, is incentive-compatible and individually-rational in the second stage.

Proof.

The agent's objective, given the proposed mechanism, is:

$$\max_{\tilde{\theta}}\int_{a_2}[u_2(q_2,\theta)-x_2]\tilde{\varphi}_2^{a_1(\theta)}(da_2|\tilde{\theta})$$

s.t.

$$a_1(\theta)=a_1(\tilde{\theta})$$

We can write the same objective function taking into account 2.P.4.

$$\max_{\tilde{\theta}}\int_{a_2}\int_{s_1}[u_2(q_2,\theta)-x_2]\varphi_2^{s_1,a_1(\theta)}(da_2|\tilde{\theta})\gamma(ds_1|\tilde{\theta})$$

Note that the agent must lie consistently, this is why $a_1(\theta)=a_1(\tilde{\theta})$, now if we use this condition in 2.P.1 and 2.P.3 we can see the agent participates and the solution to the problem is $\tilde{\theta}=\theta$. □

Proposition 2.P.6: The mechanism $\tilde{\varphi}_1$ with $\tilde{\varphi}_2^a$ is incentive-compatible and individually-rational in the first stage.

Proof.

The agent's objective, given the proposed mechanisms, is:

$$\max_{\tilde{\theta},\hat{\theta}}u_1(q_1(\tilde{\theta}),\theta)-x_1(\tilde{\theta})+\int_{a_2}[u_2(q_2,\theta)-x_2]\tilde{\varphi}_2^{a_1(\tilde{\theta})}(da_2|\hat{\theta})$$

s.t.

$$a_1(\hat{\theta})=a_1(\tilde{\theta})$$

It is straightforward the previous problem can be written as one variable maximization problem because of $a_1(\hat{\theta}) = a_1(\tilde{\theta})$. So, without loss of generality, the agent's objective is:

$$\max_{\tilde{\theta}} u_1(q_1(\tilde{\theta}),\theta) - x_1(\tilde{\theta}) + \int_{a_2} [u_2(q_2,\theta) - x_2]\tilde{\varphi}_2^{a_1(\tilde{\theta})}(da_2|\tilde{\theta})$$

The solution to this problem is $\tilde{\theta} = \theta$ because of 2.P.1 and the agent willingly participates because of 2.P.2. □

Proposition 2.P.7: The firm in the second stage willingly implements $\tilde{\varphi}_2^a$ (sequentially rational).

Proof.

Take w^* as the best profit the firm can make considering only the information disclosed by a. Then it must be true that:

$$\mathbb{E}_{\theta\sim\mu_2(a_1,a)}\left[\int_{a_2} [x_2 - c_2(q_2)]\tilde{\varphi}_2^a(da_2|\theta)\right] \leqslant w^*$$

We also know the original mechanism reacts to additional information.

$$\mathbb{E}_{\theta\sim\mu_2(a_1,a)}\left[\int_{s_1}\left(\int_{a_2} [x_2 - c_2(q_2)]\varphi_2^{s_1,a}(da_2|\theta)\right)\gamma(ds_1|\theta)\right] \geqslant w^*$$

Finally.

$$w^* \leqslant \mathbb{E}_{\theta\sim\mu_2(a_1,a)}\left[\int_{s_1}\left(\int_{a_2} [x_2 - c_2(q_2)]\varphi_2^{s_1,a}(da_2|\theta)\right)\gamma(ds_1|\theta)\right] = \mathbb{E}_{\theta\sim\mu_2(a_1,a)}\left[\int_{a_2} [x_2 - c_2(q_2)]\tilde{\varphi}_2^a(da_2|\theta)\right] \leqslant w^*$$

We can conclude the principal is already doing the best he can do. □

So far, we know the proposed mechanism is incentive-compatible because of 2.P.5 and 2.P.6, moreover it is sequentially rational as we can see in 2.P.7. Now, we will call v^* the maximum profit the principal can make using mechanisms of the form $\varphi_1 : \Theta \to A$ and $\varphi_2 : \Theta \to \Delta(A)$:

$$\mathbb{E}_\theta\left[x_1(\theta) - c_1(q_1(\theta)) + \int_{a_2} [x_2 - c_2(q_2)]\tilde{\varphi}_2^{a_1(\theta)}(da_2|\theta)\right] \leqslant v^*$$

Using the same logic used by 2.P.7, the optimal mechanism is the best the firm can do using mechanisms of the form $\varphi_1 : \Theta \to A \times \Delta(S)$ and $\varphi_2 : H^2 \times \Theta \to \Delta(A)$.

$$\mathbb{E}_\theta\left[x_1(\theta) - c_1(q_1(\theta)) + \int_{s_1}\left(\int_{a_2} [x_2 - c_2(q_2)]\varphi_2^{s_1,a_1(\theta)}(da_2|\theta)\right)\gamma(ds_1|\theta)\right] \geqslant v^*$$

Finally.

$$v^* \leqslant \mathbb{E}_\theta\left[x_1(\theta) - c_1(q_1(\theta)) + \int_{s_1}\left(\int_{a_2} [x_2 - c_2(q_2)]\varphi_2^{s_1,a_1(\theta)}(da_2|\theta)\right)\gamma(ds_1|\theta)\right] = \mathbb{E}_\theta\left[x_1(\theta) - c_1(q_1(\theta)) + \int_{a_2} [x_2 - c_2(q_2)]\tilde{\varphi}_2^{a_1(\theta)}(da_2|\theta)\right] \leqslant v^*$$

We conclude the principal can do as much using mechanisms with signals as he can do using mechanisms without them. There is no way to improve the output of the game by disclosing additional information besides the allocation. ■

5. Extensions and future work

5.1. The only observable part of the allocation is q_1.

Consider a scenario where only the items purchased by any agent are observable, while the exact amount they paid remains undisclosed. This situation is commonplace, as we often encounter goods and services in the possession of others without having precise knowledge of their prices. In this extended context, our focus narrows to the observability of only the quantity, denoted as q, while the corresponding price charged by the principal, denoted as x, remains unobservable. It's important to note that any allocation is represented as a tuple (q,x), and in this context, we limit our observation to just one component of the tuple.

5.1.1. Additional assumptions

1. Histories only capture q_1 instead of a_1.
   The new set of relevant histories has the next form.
   - $h^1 = \varnothing$
   - $h^2 = (\mathbb{M}_1, \pi_1, s_1, q_1)$
   - $h^3 = ((\mathbb{M}_1, \pi_1, s_1, q_1), (\mathbb{M}_2, \pi_2, s_2, q_2))$

   Where we can see that only q is disclosed and x is never seen, except by the agent.
2. Separability utility functions.
   The agent's utility function can be written with transfers that impact only linearly.
   $$U(a_1, a_2, \theta) = u_1(q_1, \theta) - x_1 + u_2(q_2, \theta) - x_2 \tag{5.1.A.1}$$
   The principal's payoff function can be written as:
   $$W(a_1, a_2, \theta) = x_1 - c_1(q_1) + x_2 - c_2(q_2) \tag{5.1.A.2}$$
3. Deterministic mechanism in the first stage
   We will work under the assumption that the mechanism in the first stage has deterministic allocation rules. So, a mechanism in the first stage is $\varphi_1 : \Theta \to A \times \Delta(S)$ and it can be written as a tuple $(a_1(\theta), \gamma(./\theta))$ where $a_1 : \Theta \to A$ and $\gamma : \Theta \to \Delta(S)$. This is because the analysis becomes relatively simpler than a scenario where any randomization is available. The assumption is not without loss of generality, but we hope it can be adapted.

5.1.2. The problem

The problem can be written exactly as the problem seen in the previous section. Taking into account mechanisms $\varphi_2 : \Theta \to \Delta(A)$, the **Second stage problem** given some h^2 is:

$$\max_{\varphi_2} \mathbb{E}_{\theta \sim \mu_2(h^2)} \left[\int_{a_2} (x_2 - c_2(q_2)) \varphi_2(da_2|\theta) \right]$$

s.t.

$$\int_{a_2} (u_2(q_2, \theta) - x_2) \varphi_2(da_2|\theta) \geqslant 0 \qquad \forall \theta \in \Theta \tag{E.II-IR}$$

$$\int_{a_2} (u_2(q_2,\theta) - x_2)\varphi_2(da_2|\theta) \geqslant \int_{a_2} (u_2(q_2,\theta) - x_2)\varphi_2(da_2|\tilde{\theta}) \qquad \forall \tilde{\theta} \neq \theta \tag{E.II-IC}$$

Let $\varphi_2 : H^2 \times \Theta \to \Delta(A)$ be a solution to the problem above. The **First stage problem** is:

$$\max_{a_1,\gamma} \mathbb{E}_\theta \left[x_1(\theta) - c_1(q_1(\theta)) + \int_{s_1} \left(\int_{a_2} [x_2 - c_2(q_2)]\varphi_2^{s_1,q_1(\theta)}(da_2|\theta) \right) \gamma(ds_1|\theta) \right]$$

s.t.

$$u_1(q_1(\theta),\theta) - x_1(\theta) + \int_{s_1} \left(\int_{a_2} [u_2(q_2,\theta) - x_2]\varphi_2^{s_1,q_1(\theta)}(da_2|\theta) \right) \gamma(ds_1|\theta) \geqslant 0 \qquad \forall \theta \in \Theta \tag{E.I-IR}$$

$$\begin{aligned} u_1(q_1(\theta),\theta) - x_1(\theta) + \int_{s_1} \left(\int_{a_2} [u_2(q_2,\theta) - x_2]\varphi_2^{s_1,q_1(\theta)}(da_2|\theta) \right) \gamma(ds_1|\theta) \geqslant \\ \int_{a_1} \int_{s_1} \mathcal{U}^{s_1,q_1}(\theta)\varphi_1(ds_1, da_1|\tilde{\theta}) \qquad \forall \tilde{\theta} \neq \theta \end{aligned} \tag{E.I-IC}$$

Suppose we try to replicate the same idea of Theorem 2 and exactly the same proof then we will fail because every agent is willing to lie in the first stage paying the lowest price for the same quality. Since this is the only part of the allocation that remains observable the reason why agents used to pay more receiving the same q is because they want a specific signal set. The proof makes this decision separable, therefore the obvious deviation would be going for the cheapest option. Even when this is true, the theorem still holds and the proof follows similar steps but the proposed mechanism charges every agent the same as long as they have the same q. If many types share the same q we charge the minimum payment among them, while the remaining part will be charged in the second stage, so basically they end up paying the same.

5.2. Nondeterministic allocations

A natural extension would be a proof for Theorem 2 when the first mechanism is not constrained to be deterministic in the allocation. So we have to consider a scenario where $\varphi_1 : \Theta \to \Delta(A \times S)$ instead of $\varphi_1 : \Theta \to A \times \Delta(S)$. Note that several parts of the proof are based on the deterministic part of the mechanism because it is easier to detect possible lies given a past allocation. However, when a set of allocations have positive probabilities the principal can not detect if the reported type is a lie in the second stage given some realization of the first mechanism, at least it becomes harder.

In the deterministic scenario if the agent has decided to report some type then he has committed to report certain types in the future that actually received exactly the same allocation as her, so she is not interested in misreporting because he was not interested in the beginning. But now, when the agent reports some type maybe she gets an allocation that has positive probabilities with types that received different allocations than her in the first stage, therefore the randomization proposed in the proof is not necessarily incentive-compatible, at least we cannot prove it.

5.3. Non-fully persistent types

The presence of non-fully persistent types could be another interesting way to take this research further. Specifically, we are talking about a scenario where the type in the second stage could be potentially different from the first stage type, they could be correlated but not perfectly. We know in some scenarios with fully persistent types the designer is not able to change the outcome and it is trivial to note the same result holds when types are completely independent of each other. But we do not know what could happen in between.

Something surprising is the fact that Doval and Skreta 2022 only considers fully persistent types and it is still used as a valid argument in Doval and Skreta 2021 even when types are not fully persistent in this scenario. A classic approach is setting $\theta_t \in \Theta_t$ where $\theta_1 \sim U[0,1]$ while $\theta_2 \in \{L, H\}$ such that $H > L > 0$. Then we set as an additional assumption that $\mathbb{P}(\theta_2 = H) = F(\theta_1)$ because it is intuitive to think the bigger the first stage type, the bigger the chances the second type is also "big", as a kind of positive correlation. This is actually the context presented in Doval and Skreta 2021 and types are clearly non-fully persistent.

6. Conclusion

This paper provides an answer to an informational problem in a dynamic mechanism design in the absence of commitment, and it opens the door to analyzing when it is relevant to include additional information or if it benefits the principal. This result can be used in a large number of classic problems when transferences only affect linearly.

The principal objective of this document is to show it is possible to expand the ideas of Dworczak 2020 in different contexts. That pure information intermediation is actually worthless and cannot be used as an additional way to extract rents from the agents. We know this has a lot of possible extensions to make the result even more robust but it is a good starting point for further research in limited commitment.

At the same time, we have shown a possible counterintuitive result saving time for future research where the scenario is included under our specifications to not explore mechanisms with an informational part. We hope to expand this research in several directions already explained so we can include even more cases in our results.

References

Bergemann, Dirk, Alessandro Bonatti, and Tan Gan (2022). "The economics of social data". In: *The RAND J. Econ.* 53.2, pp. 263–296.

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris (2015). "The limits of price discrimination". In: *Am. Econ. Rev.* 105.3, pp. 921–957.

Bester, Helmut and Roland Strausz (2001). "Contracting with imperfect commitment and the revelation principle: the single agent case". In: *Econometrica* 69.4, pp. 1077–1098.

Doval, Laura and Vasiliki Skreta (2021). "Purchase history and product personalization". In: *arXiv preprint arXiv:2103.11504*.

– (2022). "Mechanism design with limited commitment". In: *Econometrica* 90.4, pp. 1463–1500.

Dworczak, Piotr (2016). "Mechanism Design with Aftermarkets: On the Impossibility of Pure Information Intermediation". In.

– (2020). "Mechanism design with aftermarkets: Cutoff mechanisms". In: *Econometrica* 88.6, pp. 2629–2661.

Dworczak, Piotr and Anton Kolotilin (2019). "The persuasion duality". In: *arXiv preprint arXiv:1910.11392*.

Dworczak, Piotr and Alessandro Pavan (2022). "Preparing for the worst but hoping for the best: Robust (Bayesian) persuasion". In: *Econometrica* 90.5, pp. 2017–2051.

Freixas, Xavier, Roger Guesnerie, and Jean Tirole (1985). "Planning under incomplete information and the ratchet effect". In: *The review economic studies* 52.2, pp. 173–191.

Fudenberg, Drew and Jean Tirole (1991). "Perfect Bayesian equilibrium and sequential equilibrium". In: *journal Econ. Theory* 53.2, pp. 236–260.

Kamenica, Emir and Matthew Gentzkow (2011). "Bayesian persuasion". In: *Am. Econ. Rev.* 101.6, pp. 2590–2615.

Matyskova, Ludmila and Alfonso Montes (2023). "Bayesian persuasion with costly information acquisition". In: *J. Econ. Theory*, p. 105678.

McAfee, R Preston and Daniel Vincent (1997). "Sequentially optimal auctions". In: *Games Econ. Behav.* 18.2, pp. 246–276.

Mussa, Michael and Sherwin Rosen (1978). "Monopoly and product quality". In: *J. Econ. theory* 18.2, pp. 301–317.

Myerson, Roger B (1981). "Optimal auction design". In: *Math. operations research* 6.1, pp. 58–73.

– (1982). "Optimal coordination mechanisms in generalized principal–agent problems". In: *J. mathematical economics* 10.1, pp. 67–81.

Nguyen, Anh and Teck Yong Tan (2021). "Bayesian persuasion with costly messages". In: *J. Econ. Theory* 193, p. 105212.

Skreta, Vasiliki (2006). "Sequentially optimal mechanisms". In: *The Rev. Econ. Stud.* 73.4, pp. 1085–1111.

A. Second Theorem in 5.1

Proof

Consider $\varphi_1(\theta) = (a_1(\theta), \gamma(.|\theta))$ and its induced mechanisms in the second stage, as a **solution** to the problem. Then it must be true that:

The whole mechanism is incentive-compatible, in particular, a constant report through stages does not represent a profitable deviation.

$$\begin{aligned} u_1(q_1(\theta),\theta) - x_1(\theta) + \int_{s_1}\left(\int_{a_2}[u_2(q_2,\theta)-x_2]\varphi_2^{s_1,q_1(\theta)}(da_2|\theta)\right)\gamma(ds_1|\theta) \geqslant \\ u_1(q_1(\tilde{\theta}),\theta) - x_1(\tilde{\theta}) + \int_{s_1}\left(\int_{a_2}[u_2(q_2,\theta)-x_2]\varphi_2^{s_1,q_1(\tilde{\theta})}(da_2|\tilde{\theta})\right)\gamma(ds_1|\tilde{\theta}) \qquad \forall\tilde{\theta}\neq\theta \end{aligned} \tag{5.1.P.1}$$

The agent must willingly participate in both stages.

$$u_1(q_1(\theta),\theta) - x_1(\theta) + \int_{s_1}\left(\int_{a_2}[u_2(q_2,\theta)-x_2]\varphi_2^{s_1,q_1(\theta)}(da_2|\theta)\right)\gamma(ds_1|\theta) \geqslant 0 \qquad \forall\tilde{\theta}\neq\theta \tag{5.1.P.2}$$

$$\int_{a_2}[u_2(q_2,\theta)-x_2]\varphi_2^{s_1,q_1(\theta)}(da_2|\theta) \geqslant 0 \qquad \forall\tilde{\theta}\neq\theta \tag{5.1.P.3}$$

Consider the next mechanisms for both stages:

1. The first stage mechanism is:

$$\tilde{q}_1(\theta) = q_1(\theta)$$

$$\tilde{x}_1(\theta) = \underline{x}_1(q_1(\theta)) = min_{\hat{\theta}\in\mathbb{C}(q_1(\theta))}(x_1(\hat{\theta}))$$

   Where $\mathbb{C}(q) = \{\theta\in\Theta / q_1(\theta) = q\}$.

2. Given some history h^2 (an allocation q previously observed), and a message $\tilde{\theta}$ in the second stage, the second stage mechanism is $\tilde{\varphi}_2^q(a_2|\theta)$.

$$\int_{x_2}\tilde{\varphi}_2^q(a_2|\theta)dx_2 = \begin{cases} \int_{x_2}\int_{s_1}\varphi_2^{q,s_1}(a_2|\theta)dx_2\gamma(ds_1|\theta) & q_1(\theta)=q \\ \int_{x_2}\Omega(a_2)dx_2 & q_1(\theta)\neq q \end{cases} \tag{5.1.P.4}$$

$$\int_{a_2}x_2\tilde{\varphi}_2^q(da_2|\theta) = \begin{cases} \int_{s_1}\int_{a_2}x_2\varphi_2^{q,s_1}(da_2|\theta)\gamma(ds_1|\theta)+(x_1(\theta)-\underline{x}_1(q)) & q_1(\theta)=q \\ 0 & q_1(\theta)\neq q \end{cases} \tag{5.1.P.5}$$

   Where $\Omega(a_2) = 1$ when $a_2 = (0,\infty)$ and $\Omega(a_2) = 0$ otherwise.

Proposition 5.1.P.6: The mechanism $\tilde{\varphi}_2^q$, given some truthful h^2, is incentive-compatible and individually-rational in the second stage.

Proof.

Given the agent has told the truth we can write 5.1.P.1 as:

$$u_1(q_1(\theta),\theta)-(\underline{x}_1(q_1(\theta))+x_1(\theta)-\underline{x}_1(q_1(\theta)))+\int_{s_1}\left(\int_{a_2}[u_2(q_2,\theta)-x_2]\varphi_2^{s_1,q_1(\theta)}(da_2|\theta)\right)\gamma(ds_1|\theta)\geqslant$$
$$u_1(q_1(\theta),\theta)-(\underline{x}_1(q_1(\theta))+x_1(\tilde{\theta})-\underline{x}_1(q_1(\theta)))+\int_{s_1}\left(\int_{a_2}[u_2(q_2,\theta)-x_2]\varphi_2^{s_1,q_1(\tilde{\theta})}(da_2|\tilde{\theta})\right)\gamma(ds_1|\tilde{\theta})\qquad\forall\tilde{\theta}\in\mathbb{C}(q(\theta))$$

Then.

$$\int_{s_1}\left(\int_{a_2}[u_2(q_2,\theta)-x_2]\varphi_2^{s_1,q_1(\theta)}(da_2|\theta)\right)\gamma(ds_1|\theta)-(x_1(\theta)-\underline{x}_1(q_1(\theta)))\geqslant$$
$$\int_{s_1}\left(\int_{a_2}[u_2(q_2,\theta)-x_2]\varphi_2^{s_1,q_1(\tilde{\theta})}(da_2|\tilde{\theta})\right)\gamma(ds_1|\tilde{\theta})-(x_1(\tilde{\theta})-\underline{x}_1(q_1(\theta)))\qquad\forall\tilde{\theta}\in\mathbb{C}(q(\theta))$$

As you can see, the proposed mechanism must be incentive-compatible since if we use 5.1.P.4 and 5.1.P.5 we will get the next expression.

$$\int_{a_2}[u_2(q_2,\theta)-x_2]\tilde{\varphi}_2^{q(\theta)}(da_2|\theta)\geqslant\int_{a_2}[u_2(q_2,\theta)-x_2]\tilde{\varphi}_2^{q(\theta)}(da_2|\tilde{\theta})$$

Moreover, the agent participates in this new mechanism in the second stage because she will receive non-negative utility from it. We can be sure of it because if no participation were a profitable deviation now, then the agent would have deviated before, by reporting the type such she pays exactly $\underline{x}_1(q_1(\theta))$ in the first stage. This is pretty intuitive since every difference from the minimum payment in the group who share q could be only explained by the induced mechanisms in the second stage, because, at the end of the day, they all receive q. □

Proposition 5.1.P.7: The mechanism $\tilde{\varphi}_1$ with $\tilde{\varphi}_2^q$ is incentive-compatible and individually-rational in the first stage.

Proof.

If we take 5.1.P.1 and rewrite it.

$$u_1(q_1(\theta),\theta)-(\underline{x}_1(q_1(\theta))+x_1(\theta)-\underline{x}_1(q_1(\theta)))+\int_{s_1}\left(\int_{a_2}[u_2(q_2,\theta)-x_2]\varphi_2^{s_1,q_1(\theta)}(da_2|\theta)\right)\gamma(ds_1|\theta)\geqslant$$
$$u_1(q_1(\tilde{\theta}),\theta)-(\underline{x}_1(q_1(\tilde{\theta}))+x_1(\tilde{\theta})-\underline{x}_1(q_1(\tilde{\theta})))+\int_{s_1}\left(\int_{a_2}[u_2(q_2,\theta)-x_2]\varphi_2^{s_1,q_1(\tilde{\theta})}(da_2|\tilde{\theta})\right)\gamma(ds_1|\tilde{\theta})\qquad\forall\tilde{\theta}\neq\theta$$

With some algebra and 5.1.P.4 together with 5.1.P.5, we get the next expression.

$$u_1(q_1(\theta),\theta)-\underline{x}_1(q_1(\theta))+\int_{a_2}[u_2(q_2,\theta)-x_2]\tilde{\varphi}_2^{q(\theta)}(da_2|\theta)\geqslant$$
$$u_1(q_1(\tilde{\theta}),\theta)-\underline{x}_1(q_1(\tilde{\theta}))+\int_{a_2}[u_2(q_2,\theta)-x_2]\tilde{\varphi}_2^{q(\tilde{\theta})}(da_2|\tilde{\theta})\qquad\forall\tilde{\theta}\neq\theta$$

So, it is straightforward that our proposed mechanism is incentive-compatible and individually rational. □

Proposition 5.1.P.8: The firm in the second stage willingly implements $\tilde{\varphi}_2^q$ (sequentially rational).

Proof.

Take w^* as the best profit the firm can make considering only the information disclosed by q. Then it must be true that:

$$\mathbb{E}_{\theta\sim\mu_2(q_1,q)}\left[\int_{a_2}[x_2-c_2(q_2)]\tilde{\varphi}_2^{q}(da_2|\theta)\right]\leqslant w^*$$

We also know the original mechanism reacts to additional information.

$$\mathbb{E}_{\theta\sim\mu_2(q_1,q)}\left[\int_{s_1}\left(\int_{a_2}[x_2-c_2(q_2)]\varphi_2^{s_1,q}(da_2|\theta)\right)\gamma(ds_1|\theta)\right]\geqslant w^*$$

Finally.

$$w^*\leqslant\mathbb{E}_{\theta\sim\mu_2(q_1,q)}\left[\int_{s_1}\left(\int_{a_2}[x_2-c_2(q_2)]\varphi_2^{s_1,q}(da_2|\theta)\right)\gamma(ds_1|\theta)\right]\leqslant \mathbb{E}_{\theta\sim\mu_2(a_1,a)}\left[\int_{a_2}[x_2-c_2(q_2)]\tilde{\varphi}_2^{q}(da_2|\theta)\right]\leqslant w^*$$

We can conclude the principal is already doing the best he can do. □

So far, we know the proposed mechanism is incentive-compatible because of 5.1.P.6 and 5.1.P.7, moreover it is sequentially rational as we can see in 5.1.P.8. Now, we will call v^* the maximum profit the principal can make using mechanisms of the form $\varphi_1:\Theta\to A$ and $\varphi_2:\Theta\to\Delta(A)$:

$$\mathbb{E}_{\theta}\left[\tilde{x}_1(\theta)-c_1(q_1(\theta))+\int_{a_2}[x_2-c_2(q_2)]\tilde{\varphi}_2^{q_1(\theta)}(da_2|\theta)\right]\leqslant v^*$$

Using the same logic used by 5.1.P.8, the optimal mechanism is the best the firm can do using mechanisms of the form $\varphi_1:\Theta\to A\times\Delta(S)$ and $\varphi_2:H^2\times\Theta\to\Delta(A)$.

$$\mathbb{E}_{\theta}\left[x_1(\theta)-c_1(q_1(\theta))+\int_{s_1}\left(\int_{a_2}[x_2-c_2(q_2)]\varphi_2^{s_1,q_1(\theta)}(da_2|\theta)\right)\gamma(ds_1|\theta)\right]\geqslant v^*$$

Finally.

$$v^*\leqslant\mathbb{E}_{\theta}\left[x_1(\theta)-c_1(q_1(\theta))+\int_{s_1}\left(\int_{a_2}[x_2-c_2(q_2)]\varphi_2^{s_1,q_1(\theta)}(da_2|\theta)\right)\gamma(ds_1|\theta)\right]= \mathbb{E}_{\theta}\left[\tilde{x}_1(\theta)-c_1(q_1(\theta))+\int_{a_2}[x_2-c_2(q_2)]\tilde{\varphi}_2^{q_1(\theta)}(da_2|\theta)\right]\leqslant v^*$$

We conclude the principal can do as much using mechanisms with signals as he can do using mechanisms without them. There is no way to improve the output of the game by disclosing additional information besides the allocation. ■