Equilibrium Rejection of a Mechanism*

Gorkem Celik and Michael Peters†

February 8, 2011

Abstract

We study a mechanism design problem in which players can take part in a mechanism to coordinate their actions in a default game. By refusing to participate in the mechanism, a player can revert to playing the default game non-cooperatively. We show with an example that some allocation rules are implementable only with mechanisms which will be rejected on the equilibrium path. In our construction, a refusal to participate conveys information about the types of the players. This information causes the default game to be played under different beliefs, and more importantly under different *higher order* beliefs, than the interim ones. We find a lower bound on *all* the implementable payoffs. We use this bound to establish a condition on the default game under which all the implementable outcomes are *truthfully* implementable, without the need to induce rejection of the mechanism.

JEL classification: C72; D82; D43

Keywords: Mechanism design; Default game; Cartel agreements

*We thank the associate editor and two referees for many useful comments and suggestions. This paper was presented at Columbia University, Universite Paris IX- Dauphine, Paris School of Economics, the University of British Columbia, University of Waterloo, University of Guelph, ESSEC, THEMA, Ecole Polytechnique, University of Washington, University of Essex, and at various conferences and workshops. We are grateful for the insightful comments we received from the audiences, and especially from Francoise Forges and Philippe Jehiel. Earlier versions of this paper benefited from detailed comments by Alberto Motta and Takuro Yamashita. We also thank the European Union (Marie Curie Reintegration Grant), SSHRC Canada, and ESSEC Research Center for financial support.

†Celik: ESSEC Business School and THEMA, Avenue Bernard Hirsch, Cergy Pontoise, 95021, FRANCE; celik@essec.fr. Peters: The University of British Columbia, #997-1873 East Mall, Vancouver BC V6T 1Z1 CANADA; peters@econ.ubc.ca.

1 Introduction

There are many mechanism design problems that involve agreements about how to play some *default game*. Cartel agreements govern how firms compete against one another; members of an auction bidding ring agree on how they should bid against each other; trade agreements limit governments' ability to use trade barriers to increase their share of trade; organizations govern the efforts of workers who might otherwise compete against one another. Binding agreements on how players should act in an otherwise strategic setting often require unanimous consent of these players. One cannot coerce a firm into a cartel or a sovereign state into an international treaty. Any potential participant can veto the agreement and revert to playing the default game non-cooperatively.

In this paper, we argue that when the outside option of the participants is the non-cooperative play of a default game, the design problem is substantially different from the standard one where the outside option is a (possibly type contingent) exogenous allocation. In particular, we show that there are allocation rules that are implementable in this setting, only if the mechanism designer offers mechanisms which will be rejected by some types of some players. Since the participation decision is type dependent, a refusal to participate conveys information that causes the default game to be played differently than it would have been if players used only their interim beliefs.

Our results provide some insight into the fact that negotiations do not always lead to successful agreements. Recent examples include the failed merger between Microsoft and Yahoo as well as Ford's refusal to fully participate in the US auto bailout program. Our approach illustrates that these failures may in fact contribute to players' objectives. In our framework, a mechanism is as much a device designed to modify outcomes when negotiations fail as it is a device that shapes an agreement.

In order to make this point, we develop an example based on a possible cartel agreement between two oligopolists, one of which has a hidden production cost.[1] If these firms cannot agree on the cartel mechanism, they will play the default Cournot game. Suppose that the designer offers an agreement which is acceptable by the low cost type of the *informed* firm but

[1] Cartel agreements by asymmetrically informed firms are already studied as examples of the design setup with default game by Cramton and Palfrey (1990 and 1995). Similar examples can be constructed with auction and public good provision games, as we have done in earlier versions of this paper.

not by its high cost type. In this case, this firm will be *signaling* its cost level to its *uninformed* rival with its decision to participate in the cartel. Accordingly, whenever the cartel agreement is rejected, these firms will play the default Cournot game under the updated beliefs (and the updated higher order beliefs) on the cost of the informed firm.

The induced modification in the information structure changes the equilibrium default game behavior of the firms involved. In particular, once the low cost type of the informed firm is given the ability to signal its cost to the rival firm, it will act more aggressively (choose a higher output level) in the default Cournot game. Therefore the initially uninformed firm will expect a lower Cournot profit in comparison to its profit under the interim beliefs. In Section 2, we show that the designer can use this particular feature of the Cournot game to implement an allocation rule which would have been unacceptable to the uninformed firm if the default game was played under the interim beliefs.

In a standard mechanism design setting, where the outside option is an exogenous allocation, any potential participant has the opportunity to reject the mechanism and get the reservation payoff she would have received from this exogenous allocation. It is incumbent upon the designer to ensure that each participant receives a payoff at least as large as this reservation payoff. In contrast, when the outside option is a default game, the example we outlined above indicates a role for the designer in the determination of the reservation payoff as well.[2]

Participation decisions convey relevant information only when these decisions are type dependent and therefore non-degenerate. This means that a manipulation of the reservation payoffs requires an *equilibrium rejection*, i.e., rejection of the mechanism on the equilibrium path with positive probability. Existence of allocation rules which are implementable only through equilibrium rejection presents a major difficulty for the characterization of *all* the implementable allocation rules. Nevertheless we show that the default game yields a lower bound on the implementable payoffs. With the help of this lower bound we find a condition on the default game under which all implementable allocation rules are *truthfully* implementable without the need to

[2] A government bailout plan is another example of a mechanism where those who choose not to participate are still affected by its parameters. Philippon and Skreta (2010) study optimal bailout mechanisms to jump-start failing markets. They show that a firm's refusal of the bailout signals its confidence in the performance of its assets without a direct help from the government. See also Tirole (forthcoming).

induce an equilibrium rejection. In Section 2, we give the intuition for these two observations by using our Cournot example. In Section 3, we develop them into general results in the framework of a model with an arbitrary default game.

When the outside option is a default game, a cartel mechanism generates a sequential game of incomplete information between the firms involved. A sensible treatment of this game demands a sequential rationality restriction, which requires the firms' actions in the default game to be consistent with the beliefs they hold at the time they act. This justifies our choice of *Perfect Bayesian equilibrium* as the solution concept. In contrast, using the Bayesian equilibrium (which is not perfect) would have eliminated the relevance of the belief updates by the firms. This is best illustrated by Myerson's model of *games with contracts* (1991, Chapter 6). In this model, a mechanism can still instruct the complying players how they should play the default game when there exists some other player(s) rejecting it. Since Myerson uses Bayesian equilibrium as the solution concept, there is no sequential rationality restriction on the post-rejection instructions of the mechanisms. If our default Cournot game is played with such *Myerson mechanisms*, in the event that one firm rejects the mechanism and the other one accepts, the mechanism may instruct the complying firm to flood the market by setting a high enough production level. Such an instruction rules out the possibility of making a profit by rejecting the mechanism and fixes the outside option for each firm as zero profit.[3] Hence, the problem reduces to a standard design problem, eliminating the need for an equilibrium rejection.[4,5]

[3] What is relevant here is not the magnitude of the rejection profit, but the fact that it is independent of the allocation rule implemented by the mechanism. Under complete information, Myerson shows that any implementable allocation rule is implementable with unanimous acceptance of a mechanism, which punishes a rejecting player by *minimaxing* his payoff in the default game.

[4] *Auctions with externalities*, as studied by Jehiel, Moldovanu, and Stacchetti (1996 and 1999) constitute another example of such mechanisms. In this setting, a bidder who does not acquire the auctioned object may incur a negative externality if a *competitor* receives the object. Jehiel, Moldovanu, and Stacchetti show that the seller may extract surplus even from the bidders who do not acquire the object. The seller achieves this by threatening the bidders to give the object to their strongest competitor if they do not participate.

[5] Another way of modeling mechanisms which are not completely void after rejection is proposed by Dequiedt (2006). According to his model, once a mechanism is rejected, players choose sequentially rational actions in the default game. However, a rejected mechanism can still send messages to players, providing relevant information on the types

When the outside option is an exogenously specified allocation rule, demanding sequential rationality off the equilibrium path does not restrain the mechanism designer. Moreover, as Myerson demonstrates, once we give up sequential rationality, a default game boils down to an allocation rule. If one is willing to use Bayesian equilibrium instead of Perfect Bayesian, there is no conceptual difference between default games and type contingent allocations as outside options. On the other hand, if the objective is to understand the restrictions that the default game imposes, then the relevant solution concept must be Perfect Bayesian equilibrium or some other refinement of Bayesian equilibrium based on sequential rationality.

2 The Example

2.1 The Cournot Game with Private Cost

We build our example on an industry with two firms which are (potential) producers of an homogenous good. Both firms have linear cost functions. Firm 1 has the unit cost 0.7. The unit cost of firm 2 is either high (h =1) or low (l =0.65). The realization of its own unit cost is private information for firm 2.[6] The inverse demand function for the good is given as $P = 1 - (q_1 + q_2)$, where P is the price and q_1, q_2 are the production levels for firms 1 and 2 respectively. These firms make their production decisions simultaneously to maximize their expected profit levels.

We let β_2 denote the probability that firm 1 assigns to the event that firm 2 has the low cost. This *belief* of firm 1 is common knowledge between the firms. In the Appendix (Section 5.1), we show that the resulting game of incomplete information has a unique *Bayesian equilibrium* (BE), where firm 1 sets $q_1 = \frac{0.6-0.35\beta_2}{4-\beta_2}$, the low cost type of firm 2 sets $q_2(l) = \frac{0.4}{4-\beta_2}$, and the high cost type of firm 2 sets $q_2(h) = 0$. The expected profit level of each firm is the square of its equilibrium output level. Notice that, on the relevant domain $[0,1]$, the expected profit of firm 1, $\pi_1(\beta_2) = \left(\frac{0.6-0.35\beta_2}{4-\beta_2}\right)^2$, is a decreasing function of the belief on firm 2's type, and the profit of the

of the complying players.

[6] Firm 1 can be thought as an incumbent, whose cost structure is already revealed by its earlier conduct in the industry. Firm 2 may be a potential entrant, whose production technology is private information.

low type of firm 2, $\pi_2(l, \beta_2) = \left(\frac{0.4}{4-\beta_2}\right)^2$, is increasing in the same parameter. In other words, firm 1 loses out as it becomes more likely that its rival has the low cost, whereas firm 2 (with low cost) benefits from being perceived as the low cost.

We construct our example under the assumption that firm 1's belief on firm 2's type is uniform at the start of their interaction, i.e., $\beta_2 = 0.5$ at the *interim stage*. However, what is central to our study is understanding how these firms change their behavior when there is a change in their beliefs. Therefore, the profit functions $\pi_1(\cdot)$ and $\pi_2(l, \cdot)$, which are derived as functions of arbitrary beliefs, will prove to be useful throughout our analysis.

2.2 The Cartel Agreement

Suppose these two firms are able to sign a cartel agreement prior to making their production decisions. Following Cramton and Palfrey (1990), we model the cartel as a mechanism that is offered by a third party, which we will call the *designer*. The designer does not know the type of firm 2, and does not posses any private information herself. Our aim in this paper is to discuss what this designer is capable of doing, rather than what she would choose to do. Therefore we will not be very specific on the designer's objective for now. She may be maximizing a weighted average of the firms' expected profits or any other function of the firms' production and profit levels.

The mechanism induces a message game, where the messages from the two firms are mapped into output levels for the firms and monetary side transfers between them.[7] Firms maximize their expected profit level net of the side transfer. When offered a mechanism, each firm has an inalienable right to reject it and play the Cournot game non-cooperatively. Following the literature, we assume that the firms make their ratification decisions simultaneously. If both firms accept the mechanism, then they send their messages to the mechanism, which in turn determines the output and side transfer levels. If either one of the firms rejects the mechanism, then they

[7] As is common in the literature on mechanism design, we assume that the rules of the mechanism are enforceable once the mechanism is accepted by the participants. An explicit cartel agreement which allows the firms to coordinate their output levels and to make side transfers to each other could be outlawed by antitrust laws. Yet, tacit agreements with side transfers disguised as unrelated legitimate payments are harder to rule out. Cramton and Palfrey (1990) provide several real life examples to cartel mechanisms which are overlooked or sometimes even encouraged by the governments.

learn which firm(s) rejected it and play the Cournot game by choosing their production levels simultaneously. The design problem in this setup is non-standard since the rejection payoffs are not exogenously specified but are determined by the subsequent actions of the firms.

2.3 The Equilibrium

After the announcement of a mechanism, the interaction between the firms can be considered as a sequential game of imperfect information. The solution concept we consider here is *Perfect Bayesian equilibrium* (PBE). This solution concept is defined as a collection of *sequentially rational strategies* (which govern the ratification decisions of the firms, their message choices if the mechanism takes effect, and their production decisions if the mechanism is rejected) and *consistent beliefs* (on the type of firm 2 after observing this firm's ratification decision).[8] We provide the formal definitions of strategies, beliefs, and PBE in Section 3 within the framework of a more general model allowing for an arbitrary default game.

An *allocation rule* in this environment is defined as a mapping from the set of the firms' type profiles (in the context of our example, this is a binary set) to randomizations over the production and side transfer levels of the firms. A mechanism *implements* an allocation rule if there exists a PBE after the announcement of the mechanism, which supports the allocation rule in question. An allocation rule is called *implementable* if there exists a mechanism implementing it.[9]

In a *direct revelation mechanism*, the message set for each firm is identical to its type space (implying a singleton message set for firm 1 and a binary one for firm 2 in our example). Suppose there exists a PBE after the announcement of a direct revelation mechanism such that all types of all firms accept the mechanism and reveal their types truthfully with their messages. In this case, the resulting allocation rule is called *truthfully implementable.*

[8] In what follows, we make our main points by studying equilibria where both ratification decisions are on the equilibrium path. Therefore utilizing an alternative solution concept which extends the consistency requirement for off the equilibrium path beliefs (such as *Sequential equilibrium*) would not enrich the discussion.

[9] The implementation concept we use is "weak" implementation. That is, a mechanism implements an allocation rule if the game induced by the mechanism has an equilibrium supporting the allocation rule (as opposed to all of its equilibria supporting the allocation rule).

2.4 Off the Equilibrium Path Beliefs

Whether a firm will accept a cartel mechanism depends on the continuation payoff it expects from accepting or rejecting it. Our discussion will be mainly based on the rejection payoff to be received from the default Cournot game. If this game is played under the interim belief $\beta_2 = 0.5$, firm 1 and the low cost type of firm 2 receive the BE profit levels of $\pi_1(0.5) = \left(\frac{17}{140}\right)^2 \approx 1.4745 \times 10^{-2}$ and $\pi_2(l, 0.5) = \left(\frac{4}{35}\right)^2 \approx 1.3061 \times 10^{-2}$ respectively. However, Perfect Bayesian equilibrium allows for updating this belief after observing the ratification decision of firm 2. For instance, if a rejection is fully attributed to its high cost type, then the low cost type of firm 2 would receive only $\pi_2(l, 0) = 1 \times 10^{-2}$ from the default Cournot game. This observation suggests that the designer may truthfully implement an allocation rule which leaves a payoff of 1×10^{-2} to this firm: If both types of firm 2 are expected to accept a direct revelation mechanism, it is possible to assign the degenerate belief of $\beta_2 = 0$ in case of an off the equilibrium path rejection by firm 2.

Manipulation of the belief on the type of a party who *unexpectedly* rejects a mechanism is well studied in the design literature. Standard solution concepts, including the Perfect Bayesian equilibrium, do not put much restriction on such off the equilibrium path beliefs. The earlier literature is mostly concerned with how to refine such beliefs to find more plausible ways to outline what is feasible in a design setup.[10] In contrast to this literature, we do not employ any such refinement here. Instead, we take almost the opposite route and examine a larger class of equilibria, where rejection of the mechanism may be on the equilibrium path.[11]

[10] This idea underlines the concepts of *durability* (Myerson and Holmstrom, 1983), *attainability* (Crawford, 1985), *resilience* (Lagunoff, 1995), and *ratifiability* (Cramton and Palfrey, 1995). Imposing ratifiability of a mechanism may rule out joint profit maximization of asymmetrically informed Cournot duopolists (Cramton and Palfrey, 1995) and efficient collusion of bidders in a second price auction with participation costs (Tan and Yilankaya, 2007). A related notion is *collusion proofness* (Laffont and Martimort, 1997, 2000, and more recently Che and Kim, 2006), which requires the mechanisms to have the property that one cannot find a collusive agreement to improve over the non-cooperative reporting to the mechanism.

[11] Motta (2010) also studies rejection on the equilibrium path (or more generally, type dependence of participation decisions) in an environment where agents can collude after accepting a mechanism. If the agents are not able to collude on their participation decisions as well, Motta shows that the designer can completely eliminate the costs due to collusion. This can be done by offering the agents a selective supervision scheme, where their participation decisions would reveal their private information.

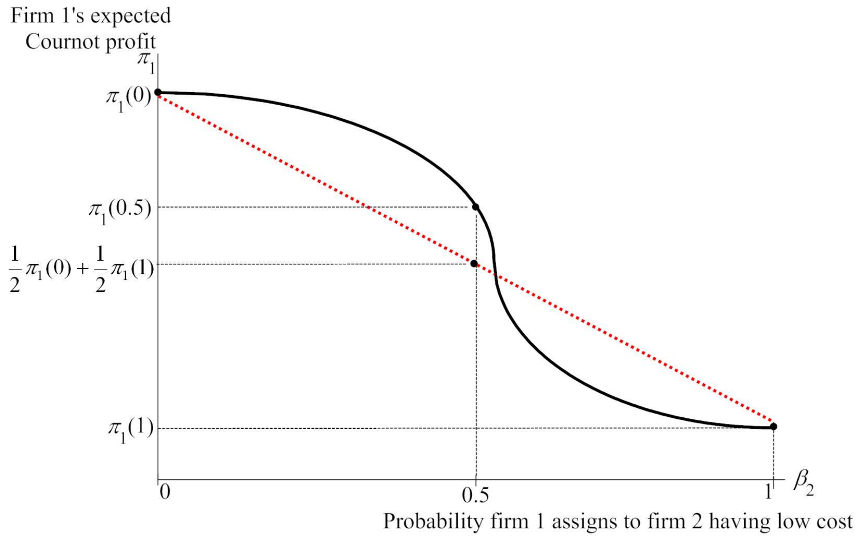
2.5 Equilibrium Path Rejection

The discussion above points out that firm 2 (with low cost) may accept a mechanism under the threat of a belief update, since its profit in the default Cournot game depends on its rival's belief β_2. We know from the construction of function $\pi_1(\beta_2)$ that the same parameter also determines firm 1's expected default Cournot profit. However, rejection of a mechanism by firm 1 cannot lead to an update of β_2, which is firm 1's own belief on the type of firm 2. In fact, in any PBE, where both types of firm 2 accept the mechanism with probability one, firm 1 receives at least its expected Cournot profit under its interim belief: $\pi_1(0.5) \approx 1.4745 \times 10^{-2}$.

We now consider another class of PBE, where firm 2's ratification behavior reveals its type. Suppose the mechanism is accepted by the low cost type of firm 2 but rejected by its high cost type. Recall that firm 1 observes firm 2's ratification decision. Therefore consistency requirement of PBE implies that firm 1 infers firm 2's type after the ratification stage. When playing the Cournot game under complete information, firm 1 receives either $\pi_1(0) = (0.15)^2$ (if firm 2 has high cost), or $\pi_1(1) = \left(\frac{1}{12}\right)^2$ (if firm 2 has low cost). The expected profit level of firm 1 is the average of these two profit levels $\frac{1}{2}\pi_1(0) + \frac{1}{2}\pi_1(1) \approx 1.4722 \times 10^{-2}$, which is smaller than its unique BE payoff from the Cournot game played under the interim belief.

The remaining task is constructing a mechanism which will indeed be accepted by the low cost type of firm 2 and rejected by its high cost type. Consider the following *simple* mechanism which does not respond to the messages by the two firms (the message set for either firm is singleton): Whenever it is accepted, the mechanism instructs the firms to set production levels $q_1 = 0$, $q_2 = 0.175$, and firm 2 to pay $\left(\frac{1}{12}\right)^2$ to firm 1 as a side payment. This corresponds to an allocation where firm 2 produces its monopoly output for its low cost type and compensates firm 1 for its foregone Cournot profit. Once this mechanism is announced, there exists a PBE of the continuation game, where firm 1 and the low cost type of firm 2 accept the mechanism, but the high cost type of firm 2 rejects it. In case of a rejection by firm 2, firm 1 learns that its rival has the high unit cost and therefore chooses its own monopoly output level of $q_1 = 0.15$. If both firms accept, then the mechanism dictates the output and side payment levels as above (The complete formal construction of the equilibrium is in the Appendix, Section 5.2). This equilibrium supports an allocation rule which leaves firm 1 with the expected profit $\approx 1.4722 \times 10^{-2}$.

Construction of this equilibrium is based on the fact that firm 1's expected profit in the default Cournot game is lower whenever firm 1 infers its competitor's type.[12] Firm 1's profit $\pi_1(\beta_2)$ is decreasing in β_2. However it is not convex. The information revealed by firm 2 allows the designer to reduce firm 1's payoff to a convex combination of the values of $\pi_1(\beta_2)$ under the two degenerate beliefs[13] (See Figure 1. For ease of demonstration, figures are not drawn to scale).



Firm 1's expected Cournot profit as a function of beliefs

Reducing firm 1's default Cournot profit by revealing more information about firm 2 may sound paradoxical. Resolution of this puzzle comes from noticing that parameter β_2 captures not only the first order belief of firm 1, but also the higher order beliefs of both firms. By providing firm 1 with information on the type of firm 2, we are also providing firm 2 with the knowledge that its rival has better information now. This affects the continuation

[12] Kim (2008) makes a similar observation in the context of common value first price auctions: If the value of the auctioned object is submodular in the bidders' signals, then a bidder prefers to be uninformed of her rival's signal.

[13] It is important to notice that, in the equilibrium we construct, firm 1 accepts the mechanism *before* observing firm 2's type dependent participation decision, when the former firm still maintains the interim belief $\beta_2 = 0.5$. However at the time of its acceptance, firm 1 is *aware* that firm 2's decision will reveal the latter firm's type and the default game would be played under one of the degenerate beliefs. For this reason, the relevant reservation payoff of firm 1 is $\frac{1}{2}\pi_1(0) + \frac{1}{2}\pi_1(1)$.

behavior of both firms in the Cournot game. In particular, when its type is known to firm 1, the low cost type of firm 2 chooses a higher output level (in comparison to its optimal output choice under the interim belief). This sequentially rational response of firm 2 to the belief update is the driving force for the reduction in firm 1's expected Cournot profit.

At first glance, it seems as if the allocation rule generated by this equilibrium can be *truthfully* implemented with the following direct revelation mechanism: Whenever firm 2 reports low cost to the mechanism, the output and side transfer levels are chosen as above ($q_1 = 0$, $q_2 = 0.175$, and firm 2 pays $\left(\frac{1}{12}\right)^2$ to firm 1). Whenever firm 2 reports high cost, the mechanism instructs the firms to mimic their non-cooperative play of the Cournot game under the belief that firm 2 has high cost ($q_1 = 0.15$, $q_2 = 0$ with no side transfer). The problem with this direct revelation mechanism is that if both types of firm 2 accept this mechanism for sure, then firm 1 would have the option of rejecting it and playing the Cournot game under the interim belief. This deviation provides firm 1 with the expected profit of $\pi_1(0.5)$, which is larger than 1.4722×10^{-2}. This observation also implies that any truthfully implementable allocation rule must leave firm 1 with a payoff at least as large as $\pi_1(0.5)$.

Our example establishes the existence of an implementable allocation rule which is not truthfully implementable. A question of interest here is whether a designer with *plausible* preferences would find such an allocation rule preferable to the truthfully implementable ones. To see the answer to this question, first notice that the industry profits are maximized with the PBE we construct above: In both states of nature, the firm with the lower cost produces its monopoly output level, and the firm with the higher cost shuts down. Moreover, firm 1's expected share of these maximized industry profits is lower than any payoff sustainable for this firm with truthful implementation. Accordingly, this allocation rule dominates all the truthfully implementable ones for a designer whose objective function is the weighted average of the firms' (ex-ante) payoffs, with the higher weight assigned to firm 2's share.

2.6 A Lower Bound on the Implementable Payoffs

The example we developed above indicates that a firm may be forced to accept a mechanism with a lower payoff, under the threat that credible information on its rival will be revealed. Since this information is revealed through the non-degenerate ratification behavior of the rival, the same payoff cannot be truthfully implemented with a direct revelation mechanism unanimously accepted by all the involved parties. Truthfully implementable allocation rules are easily identified with two sets of constraints, ensuring the participation of all players and truthful revelation of their private information. However, existence of implementable but not truthfully implementable allocation rules complicates the characterization of what is implementable. In this part of the paper, we first find a lower bound on firm 1's payoff for *all* the implementable allocation rules, including the ones feasible only through equilibrium rejection. Then we use this lower bound to derive a sufficient condition (over the parametrization of our Cournot example) under which all the implementable allocation rules are also truthfully implementable.

We start with showing that the designer can reduce firm 1's payoff even further by inducing firm 2 to reveal only *partial* information on its type. To see this, consider a mechanism which makes the high type of firm 2 indifferent between accepting or rejecting. Suppose the low cost type of firm 2 accepts this mechanism with probability one, but the high cost type accepts it with probability 1/4 only. Now consider the updated beliefs of firm 1 after observing its rival's ratification decision. If firm 1 observes a rejection (which happens with probability $(1/2) \times (3/4) = 3/8$), it believes that its rival has the high cost for sure ($\beta_2 = 0$). Otherwise, when firm 1 observes an acceptance (with probability 5/8), the Bayes formula reveals the conditional probability of facing a low cost rival as

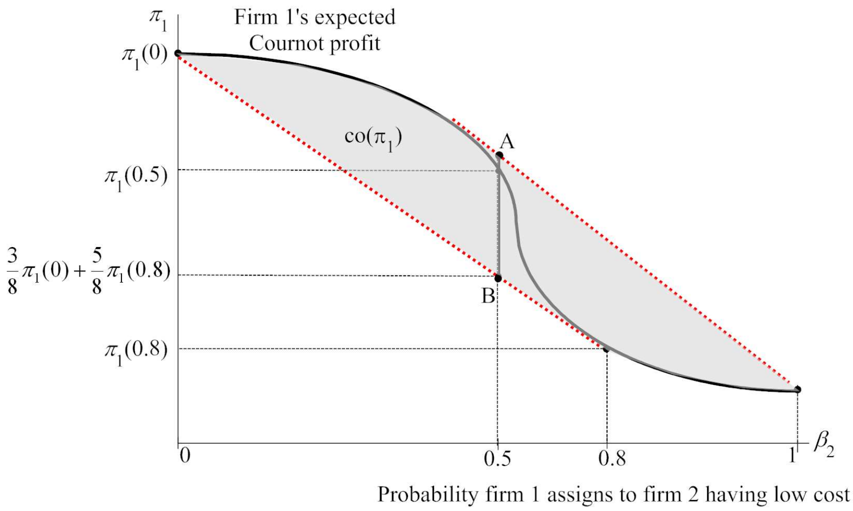
$$\beta_2 = \frac{1/2}{(1/2) + (1/2) \times (1/4)} = 0.8.$$

By rejecting the mechanism, firm 1 guarantees to play the default Cournot game under either one of these two *posterior* beliefs. This provides firm 1 with an expected Cournot profit of $\frac{3}{8}\pi_1(0) + \frac{5}{8}\pi_1(0.8) \approx 1.4688 \times 10^{-2}$, which is lower than the expected profit when the rival's type is fully revealed.

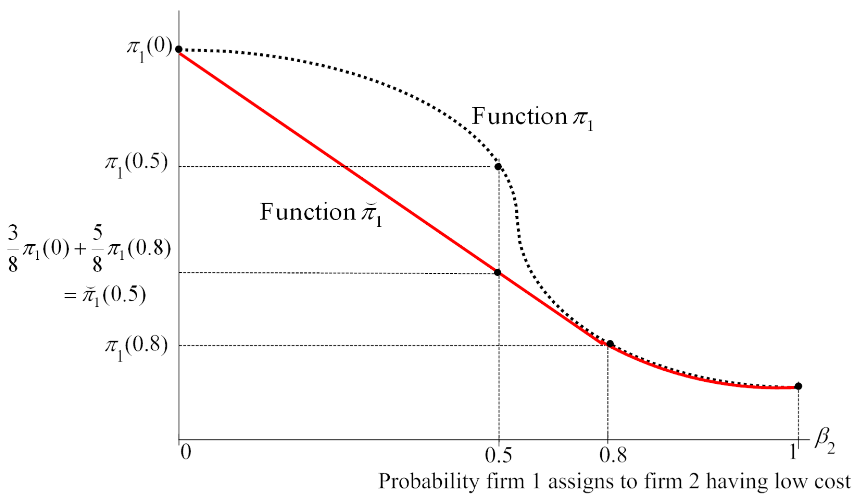
The ratification probabilities above are not chosen at random. As depicted in Figures 2 and 3, firm 1's expected Cournot profit here is the value corresponding to the interim belief $\beta_2 = 0.5$ on the *biconjugate* function which borders $co(\pi_1)$ from below, where $co(\pi_1)$ is the convex hull of the graph of the function $\pi_1(\cdot)$. The biconjugate function is formally defined as

$$\breve{\pi}_1(\beta_2) = \min\{x : (\beta_2, x) \in co(\pi_1)\}.$$

Notice that $\breve{\pi}_1(\cdot)$ is the largest convex function that is weakly smaller than $\pi_1(\cdot)$ for every value of β_2.



Convex hull of the profit function



Derivation of the biconjugate function

Whatever mechanism is offered by the designer, firm 1 always has the option of rejecting it and triggering the default Cournot game. By doing so, firm 1 receives a Cournot profit on function $\pi_1(\cdot)$. The exact level of the profit will be determined by what firm 1 learns from the rival's ratification behavior. After observing this behavior, firm 1's belief will be updated to either one of the two posterior beliefs. The values of the posterior beliefs depend on the nature of the (randomized) ratification decisions of the types of firm 2. However the Bayes rule requires that the *expected* posterior equals the interim belief $\beta_2 = 0.5$. At the ratification stage, what is relevant for firm 1 is the expected value of the two profit levels corresponding to these two posteriors. Depending on firm 2's ratification strategy, this expected profit can be anywhere on the line segment $[AB]$ drawn in Figure 2. Since the expected Cournot profit is at least as large as $\breve{\pi}_1(0.5) \approx 1.4688 \times 10^{-2}$, this value constitutes a lower bound on the implementable payoffs for firm 1.

The way that we make use of the biconjugate of function $\pi_1(\cdot)$ closely resembles Kamenica and Gentzkow's (forthcoming) utilization of the *concave closure* of a *sender*'s payoff as a function of a *receiver*'s posterior in a persuasion game environment. In their formulation, the sender chooses the posteriors of the receiver on condition that the expectation over the posteriors equals the prior. In this setting, the maximized value of the sender's *expected* payoff is on the concave closure of her payoff function, which borders the convex hull of the payoff function's graph from *above*.[14] In contrast to Kamenica and Gentzkow's sender - receiver setting, we study a mechanism design environment and try to outline the implementable allocation rules. This problem induces the minimization of the outside option for the participants of the mechanism. That is why we are interested in the biconjugate function, which borders the convex hull of the graph of $\pi_1(\cdot)$ from below.

The relation between function π_1 and its biconjugate will also produce a sufficient condition to rule out the need for an equilibrium rejection as part of an implementation. We have already seen that any truthfully implementable allocation rule would provide firm 1 with a payoff at least as large as $\pi_1(0.5)$, its Cournot profit level corresponding to the interim belief. If this profit level was already on the biconjugate of function π_1, that is if functions π_1 and $\breve{\pi}_1$ gave the same value for the interim belief, then revealing further information about the rival firm would not have reduced firm 1's Cournot profit. For instance, if the interim belief of firm 1 was weakly larger than 0.8

[14] Aumann and Maschler (1995) and Goemans and Fay (2009) refer to a similar concave closure to study maximization problems where the choice variable is the distribution under the constraint that its expectation equals a constant.

(instead of being equal to 0.5) in our Cournot example, then we would not be able to enlarge the implementable set beyond the truthfully implementable allocation rules.

3 The Model

In this section, we study a general mechanism design setting, with an arbitrary number of agents and an arbitrary default game as the outside option. To be consistent with the language of the previous section, we continue referring to the agents of our model as *firms*. $N = \{1, 2, ..., |N|\}$ is the finite set which accepts these firms as its elements. For each firm $i \in N$, we let Θ_i denote the set of finitely many types available for this firm.

At the start of the interaction, firm i observes its own type $\theta_i \in \Theta_i$, which is a random variable for the other firms. We assume that types of different firms are statistically independent. The initial beliefs of the other firms on the type of firm i are represented by the interim distribution $\beta_i^0 \in \Delta\Theta_i$. The value of $\beta_i^0(\theta_i)$ gives the probability that firms other than firm i attribute to the event that firm i has type θ_i at the time that the interaction between the firms starts.

3.1 The Default Game under Arbitrary Beliefs

In the default game, each firm i can choose an output level from the set of available output levels Q_i (or more generally an action from the set of available actions). Firm i's *direct profit* in the default game is a function of the profile of the chosen output levels and its type:

$$u_i : \times_{i \in N} Q_i \times \Theta_i \to \mathcal{R},$$

where $\mathcal{R}$ is the set of real numbers.

The set of firms, the type spaces, the available output levels, and the profit functions define a Bayesian game together with the interim beliefs specified earlier. However, since we will allow for the default game to be played under updated beliefs, we study this game under an arbitrary *belief system* $\beta = \{\beta_i\}_{i \in N}$ rather than the interim beliefs. As in the definition of the interim beliefs, β_i is an element of $\Delta\Theta_i$ and $\beta_i(\theta_i)$ yields the probability that the other firms are attributing to the event that firm i has type θ_i when they are playing the default game.

$N, \{\Theta_i\}_{i\in N}, \{Q_i\}_{i\in N}, \{u_i\}_{i\in N}$, and β constitute a Bayesian Default Game. The choice variable for firm i is its (possibly randomized) output level, which can be conditioned on the realization of its type:

$$q_i : \Theta_i \to \Delta Q_i.$$

In equilibrium, each type of each firm chooses its output level to maximize its expected direct profit in the default game. The beliefs enter into the picture in the calculation of these expected profits. Under the arbitrary belief system β, a Bayesian equilibrium of the default game is defined as a collection of output functions $\{q_i\}_{i\in N}$ such that output level (or any output level in the support of randomization) $q_i(\theta_i)$ is a solution to

$$\max_{\hat{q}_i \in Q_i} E_{\theta_{-i}} \left[u_i \left(\hat{q}_i, q_{-i}(\theta_{-i}), \theta_i \right) | \beta_{-i} \right]$$

for all $i \in N$ and all $\theta_i \in \Theta_i$.[15]

We restrict attention to default games which have at least one Bayesian equilibrium for all possible beliefs (Existence is assured whenever all Q_i's are finite sets). For equilibrium $\{q_i\}_{i\in N}$ under belief system β, expected profit for type θ_i is $E_{\theta_{-i}} \left[u_i \left(q_i(\theta_i), q_{-i}(\theta_{-i}), \theta_i \right) | \beta_{-i} \right]$. Due to the possibility of multiple equilibria, there may be more than one expected profit level consistent with some belief system. We define π_i as the correspondence that maps the type of firm i and the belief system to the expected equilibrium profits. That is, $\pi_i \left(\theta_i, \beta_i, \beta_{-i} \right)$ is the set of the expected equilibrium profit levels for type θ_i, when the default game is played under belief system $\left(\beta_i, \beta_{-i} \right)$.

As was the case with the Cournot example of the previous section, the largest function which is convex in β_{-i} and which takes values weakly smaller than the values of correspondence $\pi_i \left(\theta_i, \beta_i, \beta_{-i} \right)$ everywhere will have an important part in our analysis of the general model. With a minor abuse of terminology, we refer to this function as the biconjugate function for correspondence π_i and define it as

$$\breve{\pi}_i \left(\theta_i, \beta_i, \beta_{-i} \right) = \inf \left\{ x : \left(\beta_{-i}, x \right) \in co\left(\pi_i \left(\theta_i, \beta_i, \cdot \right) \right) \right\},$$

where $co\left(\pi_i \left(\theta_i, \beta_i, \cdot \right) \right)$ refers to the convex hull of the graph of correspondence $\pi_i \left(\theta_i, \beta_i, \cdot \right)$ for fixed values of θ_i and β_i.

[15] As is standard, subscript $-i$ refers to the collection of one variable for each firm other than firm i.

3.2 The Message Game under Arbitrary Beliefs

The cartel mechanism, if accepted, instructs the firms what output levels they should choose in the default game. Moreover, the mechanism allows for transferable payoffs between the firms by stipulating monetary side transfers. The mechanism can condition these choices on revelations by (the messages of) the firms. Formally a mechanism is comprised of a message set M_i for each firm i, an output function determining the firms' (possibly randomized) production levels $\chi : \times_{i \in N} M_i \rightarrow \Delta \times_{i \in N} Q_i$, and a side transfer function $\tau : \times_{i \in N} M_i \rightarrow \Delta \mathcal{R}^{|N|}$ subject to the constraint that the transfers the firms receive from each other add up to 0 under all states.[16]

We can construct the payoff (net of the side transfer) of a firm taking part in this mechanism by using the profit functions in the default game. Suppose that the firms send the message profile $m \in \times_{i \in N} M_i$ to the mechanism. In this case, firm i with type θ_i will end up with payoff

$$w_i(m, \theta_i) = u_i(\chi(m), \theta_i) + \tau_i(m),$$

where $\tau_i(m)$ is the transfer firm i receives (the i^{th} component of vector $\tau(m)$). Once these payoff functions are defined, an accepted mechanism induces a message game between the firms. The choice variable for each firm is its (possibly randomized) message, which can be conditioned on the realization of its type:

$$m_i : \Theta_i \rightarrow \Delta M_i.$$

Each type of each firm chooses a message to maximize its expected payoff. As in our discussion of the default game, the expected payoff is well defined only under some belief system β. Under the arbitrary belief system β, a Bayesian equilibrium of the message game is a collection of message functions $\{m_i\}_{i \in N}$ such that message (or any message in the support of randomization) $m_i(\theta_i)$ is a solution to

$$\max_{\hat{m}_i \in M_i} E_{\theta_{-i}} \left[w_i\left(\hat{m}_i, m_{-i}(\theta_{-i}), \theta_i\right) | \beta_{-i} \right]$$

for all $i \in N$ and all $\theta_i \in \Theta_i$.

[16] The model can be easily extended to allow for correlation between the randomizations over production and side transfer levels.

3.3 Ratification

After the firms observe the default game and the mechanism, they simultaneously decide whether to accept the mechanism or not. We represent firm i's ratification decision with the binary variable r_i (equals to y if firm i accepts, to n otherwise). Firm i can condition the probability of acceptance on its private information:

$$\sigma_i : \Theta_i \rightarrow [0,1].$$

We refer to $r = \{r_i\}_{i \in N}$ as a *ratification profile.* Since there are a total of $|N|$ firms and each of them can choose one of the two decisions, there are $2^{|N|}$ different possible ratification profiles. After the ratification stage, firms observe each others' ratification decisions. In other words, the realized vector r becomes public information. This observation gives the firms the opportunity to update beliefs on each other. Whenever firm i accepts the mechanism, the belief of the other firms on the type of firm i is represented by $\beta_i^y \in \Delta\Theta_i$. Similarly, $\beta_i^n \in \Delta\Theta_i$ gives the rival firms' belief if firm i rejects the mechanism.[17]

The mechanism comes into effect in the event that it is accepted unanimously by all firms (that is, when the ratification profile is a vector composed of y's). In this case, each firm sends its message to the mechanism. By using these messages as inputs, functions χ and τ determine the output and side transfer levels, and eventually the payoffs of the firms.

On the other hand, the mechanism is vetoed whenever there exists some firm(s) rejecting it at the ratification stage. Notice that there are $2^{|N|} - 1$ different ratification profiles under which the mechanism will be rejected by at least one of the firms. We let V denote the set of these $2^{|N|} - 1$ ratification profiles that lead to a veto of the mechanism. After learning the realized ratification profile r, each firm decides on its default game output level. The direct profits of the firms from the default game are determined by these output levels.

[17] With this notation, we impose the restriction that two different firms, which have observed the same off the equilibrium path decision by a rival firm, will hold the same belief about this rival. This condition is automatically satisfied when the solution concept is Sequential equilibrium. Making this restriction is also commonplace for Perfect Bayesian equilibrium.

3.4 The Equilibrium

Given the default game, the interim beliefs, and the proposed mechanism, the resulting interaction between the firms can be thought as a sequential game. We now specify the Bayesian (behavior) strategies and beliefs constituting a Perfect Bayesian equilibrium (PBE) of this game:

Definition 1 *A Perfect Bayesian equilibrium is a collection of strategies and beliefs* $\left\{\sigma_i^*, m_i^*, \{q_i^{r*}\}_{r\in V}, \beta_i^{y*}, \beta_i^{n*}\right\}_{i\in N}$ *which together satisfy the conditions listed below:*

i) $\{m_i^*\}_{i\in N}$ *constitutes a Bayesian Equilibrium of the message game induced by the mechanism under the belief system* $\{\beta_i^{y*}\}_{i\in N}$.

ii) For each ratification profile $r \in V$ *leading to a veto of the mechanism,* $\{q_i^{r*}\}_{i\in N}$ *constitutes a Bayesian Equilibrium of the default game under the belief system* $\{\beta_i^{r_i*}\}_{i\in N}$.

iii) For each firm $i \in N$ *and each type* $\theta_i \in \Theta_i$, $\sigma_i^*(\theta_i)$ *maximizes the expected continuation payoff, given the rival firms' ratification behavior* σ_{-i}^* *and the continuation strategies of all firms* $\{m_i^*\}_{i\in N}, \{q_i^*\}_{i\in N}$.

iv) $\{\beta_i^{y*}\}_{i\in N}$ *and* $\{\beta_i^{n*}\}_{i\in N}$ *are derived by the Bayes formula on the equilibrium path. That is,* $\beta_i^{y*}(\theta_i) = \frac{\beta_i^0(\theta_i)\sigma_i^*(\theta_i)}{\sum_{\hat{\theta}_i\in\Theta_i}\beta_i^0(\hat{\theta}_i)\sigma_i^*(\hat{\theta}_i)}$ *if* σ_i^* *is not constant at* 0 *and* $\beta_i^{n*}(\theta_i) = \frac{\beta_i^0(\theta_i)\left[1-\sigma_i^*(\theta_i)\right]}{\sum_{\hat{\theta}_i\in\Theta_i}\beta_i^0(\hat{\theta}_i)\left[1-\sigma_i^*(\hat{\theta}_i)\right]}$ *if* σ_i^* *is not constant at 1.*

Now that we are equipped with the formal statement of our solution concept, we can use the definitions in Section 2 to refer to the implementable and truthfully implementable allocation rules in this general setting.

3.5 The Analysis

By rejecting a mechanism, firm i guarantees that the continuation game will be the default game instead of the message game stipulated by the mechanism. Nevertheless, at the time of its rejection, firm i does not necessarily know the beliefs under which the default game will be played. The belief on firm i's type itself, β_i^{n*}, is pinned down by the equilibrium. However, the belief on the type of a rival firm, say firm j, will depend on whether firm j accepts the mechanism (β_j^{y*}) or joins firm i in rejecting it (β_j^{n*}). The realization of firm j's ratification decision is unknown to firm i at the ratification stage. What is important to notice here is that, since beliefs β_j^{y*} and β_j^{n*} are both derived from the Bayes rule (whenever firm j's ratification decision is non-degenerate), the *expected* belief on firm j's type is equal to the interim belief:

$$\left[\sum_{\theta_j \in \Theta_j} \beta_j^0(\theta_j)\,\sigma_j^*(\theta_j)\right]\beta_j^{y*} + \left[\sum_{\theta_j \in \Theta_j} \beta_j^0(\theta_j)\left(1-\sigma_j^*(\theta_j)\right)\right]\beta_j^{n*} = \beta_j^0.$$

Recall that π_i is defined as the correspondence which gives the possible profit levels when the default game is played under different beliefs. Accordingly, $\pi_i\left(\theta_i, \beta_i^{n*}, \beta_{-i}^{r_{-i}*}\right)$ gives the possible profit levels for type θ_i of firm i if the default game is played under belief system $\left(\beta_i^{n*}, \beta_{-i}^{r_{-i}*}\right)$. As we have seen in the previous section, when the default game may be played under a variety of belief systems, firm i's *expected* profit can be strictly lower than the values of π_i under the interim beliefs. We use the biconjugate function $\breve{\pi}_i$ to establish a lower bound on the equilibrium payoff of firm i.

Proposition 1 *Consider a mechanism and a PBE* $\left\{\sigma_i^*, m_i^*, \{q_i^{r*}\}_{r\in V}, \beta_i^{y*}, \beta_i^{n*}\right\}_{i\in N}$ *of the continuation game. Under this equilibrium, the expected payoff of firm i with type θ_i is at least as large as* $\breve{\pi}_i\left(\theta_i, \beta_i^{n*}, \beta_{-i}^0\right)$.

The proof is relegated to the Appendix. This result yields a lower bound on the equilibrium payoff as a function of the rejection beliefs specified by the equilibrium. In order to get a bound that refers only to the primitives of the problem, it suffices to minimize $\breve{\pi}_i$ over the beliefs on firm i.

Corollary 1 *For any implementable allocation rule, the expected payoff of firm i with type θ_i is at least as large as* $\min_{\beta_i} \breve{\pi}_i\left(\theta_i, \beta_i, \beta_{-i}^0\right)$.

The Cournot default game we covered in Section 2 is an example of the case where the biconjugate function $\breve{\pi}_i$ lies strictly below the correspondence π_i for some beliefs. In the analysis of this example we have seen that this situation brings the opportunity of reducing the payoff of a firm below its default game Bayesian equilibrium profit under the interim beliefs. However supporting any such payoff requires the construction of an equilibrium where some rival firm signals part of its private information with its ratification of the mechanism. For this signal to have an informative value, both accepting and rejecting the mechanism must be equilibrium behavior for this rival firm. Since truthful implementation demands for a unanimous acceptance of a direct revelation mechanism by all types of all firms, no such payoff is truthfully implementable.

On the other hand, if the values of function $\breve{\pi}_i$ and correspondence π_i coincide for the interim belief β^0_{-i} for all types of all firms regardless of the belief β_i, then considering rejections on the equilibrium path does not extend the set of implementable outcomes. In other words, this condition rules out the implementable allocation rules which are not truthfully implementable. We conclude our analysis with the formalization of this result, which we prove in the Appendix.

Proposition 2 *Suppose that for each firm $i \in N$, each type $\theta_i \in \Theta_i$, and each belief $\beta_i \in \Delta\Theta_i$, correspondence $\pi_i\left(\theta_i, \beta_i, \beta^0_{-i}\right)$ has a single value which equals to $\breve{\pi}_i\left(\theta_i, \beta_i, \beta^0_{-i}\right)$. Then any implementable allocation rule is also truthfully implementable.*

4 Conclusion

In this paper, we studied a mechanism design problem where players either accept a mechanism or play a default game non-cooperatively. Default games are more difficult to handle than (possibly type contingent) exogenous allocations as outside options of mechanisms. The difficulty arises from the existence of allocation rules which are implementable only if a mechanism is rejected on the path of play. Although our modeling of this problem is in line with most of the earlier literature, it is certainly not the unique way to address a mechanism design setting. We conclude the paper with a discussion of alternative modeling assumptions.[18]

Pre-play communication in the default game:

Rejection of a mechanism on the equilibrium path is crucial for implementation of certain allocation rules since a rejection has the potential to reveal information on the type of the rejecting player. In this case the default game is played in light of this additional information. An alternative way of providing players with the opportunity to signal their types is allowing for pre-play communication in the default game.[19] However exchanging *cheap*

[18]For the sake of brevity, we do not provide the technical details of the arguments we make here. A formal analysis of each alternative assumption is available from the authors.

[19]Matthews and Postlewaite (1989), Palfrey and Srivastava (1991), and Forges (1999) show that pre-play communication between players of a game dramatically extends the set of equilibria.

talk messages in a pre-play communication stage is not a perfect substitute for equilibrium rejection of a mechanism. The latter form of communication determines whether a mechanism will take effect or not, and therefore carries an inherent cost for the players, unlike sending cheap talk messages.

Unobservable ratification decisions:

We assumed in our analysis that each player's acceptance or rejection of the mechanism is observed by all the others. An alternative approach is assuming that the players only find out whether a mechanism takes effect or not, instead of learning about every individual ratification decision. Notice that, under this alternative assumption, a player who accepts the mechanism will still infer some information about its rivals by simply observing if the mechanism is unanimously accepted. In this case, rejection of the mechanism signals the existence of at least one player who has refused to participate, giving the opportunity for a belief update. Therefore, even when the individual ratification decisions are unobserved, it is still possible to construct allocation rules which are implementable only with mechanisms rejected on the equilibrium path.

Mechanisms offered by players:

Suppose the mechanism is offered not by a designer, but by one of the players who has private information. This assumption creates an *informed principal* setting.[20] The analysis we provide in this paper suggests that there are allocation rules which are implementable only if this informed principal signals her type with the choice of the mechanism. This is in contrast with the *inscrutability principle*, which applies to settings with exogenous outside options and which indicates that any available allocation rule can be supported with an equilibrium where the principal offers the same mechanism regardless of her type.

5 Appendix

5.1 BE of the Default Cournot Game

The Cournot game with private information, which we utilized to construct our example in Section 2, can be analyzed within the framework developed in Section 3. Any nonnegative output level is available to either firm, which

[20] As studied by Myerson (1983) and Maskin and Tirole (1990, 1992) for the case where the outside option is an exogenous allocation rule.

means that $Q_i = \mathcal{R}_+$ for $i = 1, 2$. The profit levels of the firms are $u_i = [1 - (q_i + q_j) - c_i]\, q_i$, where c_i indicates the unit cost. The unit cost of firm 1 equals to 0.7 and firm 2's cost level is either $l = 0.65$ or $h = 1$. Since the type space of firm 2 is binary, the belief on its type (β_2) can be represented by the probability of this firm assuming the low cost type l. The interim belief on firm 2's type is given as $\beta_2^0 = 0.5$.

When applying the definition of the Bayesian equilibrium to the Cournot game, the first point to note is the dominant strategy of firm 2 with the high cost level. Whenever firm 2 has unit cost 1, its profit function is given as $-(q_1 + q_2(h))q_2(h)$, which is maximized with the output choice $q_2(h) = 0$. Therefore the equilibrium output level of the high cost type of firm 2 is determined as zero regardless of the beliefs. Zero output brings zero profit to this firm.

Now we move to the output levels of firm 1 as well as the low cost type of firm 2. Since the latter firm has unit cost 0.65, its profit level is given by the function $(1 - q_1 - q_2(l) - 0.65)\, q_2(l)$, which is maximized with the output choice $q_2(l) = \frac{0.35 - q_1}{2}$ for the relevant values of q_1. To derive a similar reaction function for firm 1, notice that the expected output level by firm 2 is $\beta_2 q_2(l)$. Since firm 1 has cost 0.7, its reaction function is written as $q_1 = \frac{0.3 - \beta_2 q_2(l)}{2}$ for the relevant values of $q_2(l)$. When we solve for the two reaction functions simultaneously, we get the *unique* BE output levels as functions of parameter β_2:

$$q_1 = \frac{0.6 - 0.35\beta_2}{4 - \beta_2},\ q_2(l) = \frac{0.4}{4 - \beta_2}.$$

After substituting these values in the (expected) profit functions, we see that the maximized levels of the profits are $(q_1)^2$ and $(q_2(l))^2$ for these two firms.

5.2 The Mechanism and PBE of the Induced Game

The mechanism constructed in Section 2 does not make use of the messages sent after acceptance. Accordingly, message sets M_1 and M_2 are both singleton. The output levels following a unanimous acceptance of the mechanism are $\chi_1 = 0$ and $\chi_2 = 0.175$. When the mechanism takes effect, firm 2 makes the monetary transfer of $\tau_1 = \left(\frac{1}{12}\right)^2$ to firm 1. In the sequential game following the announcement of this mechanism, there exists a PBE depicted as below following the notation developed in Section 3:

- $\sigma_1^* = \sigma_2^*(l) = 1$ and $\sigma_2^*(h) = 0$,

- if $r_1 = n$ and $r_2 = y$, then $\beta_2^{r*} = 1$, $q_1^{r*} = \frac{1}{12}$, $q_2^{r*}(l) = \frac{2}{15}$, and $q_2^{r*}(h) = 0$,

- if $r_2 = n$, then $\beta_2^{r*} = 0$, $q_1^{r*} = 0.15$, $q_2^{r*}(l) = 0.1$, and $q_2^{r*}(h) = 0$ regardless of the value of r_1.

To prove that this is indeed an equilibrium, we need to establish that the strategies and the beliefs above satisfy the four conditions of PBE:

i) Since the induced message game is a degenerate game without any message choices for the firms, this condition is trivially satisfied.

ii) For $r_2 = y$, the belief $\beta_2^{r*} = 1$ implies that q_1^{r*} and $q_2^{r*}(l)$ are equal to the unique *complete information* Cournot competition output levels ($\frac{1}{12}$ and $\frac{2}{15}$) when firms have cost levels 0.7 and 0.65. Moreover, $q_2^{r*}(h) = 0$ is the dominant output level for firm 2 with high cost.

For $r_2 = n$, the belief $\beta_2^{r*} = 0$ dictates that firm 1 produces the monopoly output level 0.15 . As a best response, firm 2 with low cost produces 0.1 and firm 2 with high cost produces 0.

iii) By rejecting the mechanism, firm 1 receives the average of $\left(\frac{1}{12}\right)^2$ and the monopoly profit $(0.15)^2$. By accepting the mechanism, firm 1 guarantees exactly the same payoff. Since firm 1 is indifferent, accepting the mechanism with probability one is an optimal ratification behavior.

Firm 2 with low cost receives $(0.175)^2 - \left(\frac{1}{12}\right)^2$ by accepting the mechanism and $(0.1)^2$ by rejecting it. Since acceptance brings a larger payoff, $\sigma_2^*(l) = 1$ is optimal. Firm 2 with high cost receives a negative payoff by accepting the mechanism and 0 by rejecting it. Therefore the optimal ratification decision induces $\sigma_2^*(h) = 0$ as well.

iv) The ratification behavior separates the two types of firm 2. Bayes rule dictates that $\beta_2^{r*} = 1$ if $r_2 = y$ and $\beta_2^{r*} = 0$ if $r_2 = n$.

5.3 Proof of Proposition 1

Suppose firm i rejects the mechanism at the ratification stage. Following this rejection, the belief on the type of firm i is updated to β_i^{n*}. There are $2^{|N|-1}$ different combinations of ratification decisions of firm i's rivals. Therefore, following the rejection of firm i, the default game will be played under one of the potentially $2^{|N|-1}$ different beliefs $\left(\beta_{-i}^{r_{-i}*}\right)$ on the types of firm i's rivals. It follows from the Bayes rule that the expected value of these $2^{|N|-1}$ different beliefs is equal to the interim belief β_{-i}^0 at the ratification stage.

In the default game, all firms choose their output levels optimally. Therefore $\pi_i\left(\theta_i, \beta_i^{n*}, \beta_{-i}^{r_{-i}*}\right)$ gives the possible profit levels for type θ_i of firm i when the default game is played under belief $\left(\beta_i^{n*}, \beta_{-i}^{r_{-i}*}\right)$. Since the expected value of $\beta_{-i}^{r_{-i}*}$ is β_{-i}^{0} at the ratification stage, the expected payoff of type θ_i of firm i from rejecting the mechanism is weakly larger than $\breve{\pi}_i\left(\theta_i, \beta_i^{n*}, \beta_{-i}^{0}\right)$. Type θ_i chooses its ratification decision to maximize the continuation payoff. Accordingly, the equilibrium payoff of this type is at least as large as the lower bound on the rejection payoff $\breve{\pi}_i\left(\theta_i, \beta_i^{n*}, \beta_{-i}^{0}\right)$.

5.4 Proof of Proposition 2

We start with an arbitrary implementable allocation rule. There exists a (possibly indirect) mechanism $\mathcal{M}^*$ and a PBE $E^* = \left\{\sigma_i^*, m_i^*, \{q_i^{r*}\}_{r\in V}, \beta_i^{y*}, \beta_i^{n*}\right\}_{i\in N}$ of the continuation game supporting this rule. Suppose the same allocation rule is now offered as a direct revelation mechanism $\mathcal{M}^d$. To prove the proposition, we have to show the existence of a continuation PBE $E^d = \left\{\sigma_i^d, m_i^d, \left\{q_i^{rd}\right\}_{r\in V}, \beta_i^{yd}, \beta_i^{nd}\right\}_{i\in N}$ such that all types of all firms accept the mechanism $\left(\sigma_i^d\left(\theta_i\right) = 1, \text{ for all } i \text{ and } \theta_i\right)$, and reveal their types truthfully $\left(m_i^d\left(\theta_i\right) = \theta_i, \text{ for all } i \text{ and } \theta_i\right)$. By construction, each type of each firm receives the same payoff under E^d as under E^*.

Equilibrium E^d instructs all parties to accept the direct revelation mechanism unanimously. It follows from the Bayes rule (satisfying condition *(iv)* of the definition of PBE) that beliefs remain the same as their interim values after observing the acceptance of a firm:

$$\beta_i^{yd} = \beta_i^0 \text{ for all } i.$$

On the other hand, there is no consistency requirement for the off the equilibrium path rejection beliefs in E^d. In our construction, we set these beliefs to be the same as the rejection beliefs of equilibrium E^*:

$$\beta_i^{nd} = \beta_i^{n*} \text{ for all } i.$$

Output functions $\left\{q_i^{rd}\right\}_{r\in V}$ specify the continuation behavior of firm i following an off the equilibrium path rejection. We set these functions to be the Bayesian equilibrium output functions of the default game under the corresponding beliefs $\left\{\beta_i^{r_i d}\right\}_{i\in N}$ (satisfying condition *(ii)* of the definition of PBE).

To complete the proof, we need to show that truthful revelation of the type and acceptance of the mechanism constitute an equilibrium together with the beliefs $\left\{\beta_i^{yd}, \beta_i^{nd}\right\}_{i\in N}$ and output functions $\left\{q_i^{rd}\right\}_{r\in V, i\in N}$ specified above, after the announcement of direct revelation mechanism $\mathcal{M}^d$.

Truthful revelation: Suppose type θ_i of firm i imitates type $\hat{\theta}_i$ in the direct revelation message game. Under the interim beliefs $\left\{\beta_i^0\right\}_{i\in N}$, this deviation will bring type θ_i the same payoff as mimicking the equilibrium strategy of type $\hat{\theta}_i$ in E^* (that is, following strategy $\sigma_i^*\left(\hat{\theta}_i\right), m_i^*\left(\hat{\theta}_i\right), \left\{q_i^{r*}\left(\hat{\theta}_i\right)\right\}_{r\in V}$) after the announcement of mechanism $\mathcal{M}^*$. Since ratification, message, and output decisions are made optimally under E^*, imitating type $\hat{\theta}_i$ does not bring a strictly higher expected payoff to type θ_i of firm i (satisfying condition *(i)* of the definition of PBE).

Acceptance of the mechanism: By unilaterally rejecting this mechanism, firm i guarantees playing the default game under beliefs $\left(\beta_i^{n*}, \beta_{-i}^0\right)$. The hypothesis of the proposition implies that, when the default game is played under these beliefs, there is a unique equilibrium profit level for each type θ_i of firm i, which equals to $\breve{\pi}_i\left(\theta_i, \beta_i^{n*}, \beta_{-i}^0\right)$. Recall that β_i^{n*} is the rejection belief under equilibrium E^* following the announcement of mechanism $\mathcal{M}^*$. It follows from Proposition 1 that the equilibrium payoff for type θ_i (under both E^* and E^d) is at least as large as the deviation payoff $\breve{\pi}_i\left(\theta_i, \beta_i^{n*}, \beta_{-i}^0\right)$ (satisfying condition *(iii)* of the definition of PBE).

References

[1] Aumann, R. J. and M. B. Maschler, 1995. "Repeated games with incomplete information" MIT Press.

[2] Che, Y. and J. Kim, 2006. "Robustly collusion-proof implementation" Econometrica, 74, 1063-1107.

[3] Cramton, P. C. and T. R. Palfrey, 1990. "Cartel enforcement with uncertainty about costs" International Economic Review, 31, 17-47.

[4] Cramton, P. C. and T. R. Palfrey, 1995. "Ratifiable mechanisms: learning from disagreement" Games and Economic Behavior, 10, 255-283.

[5] Crawford, V. P., 1985. "Efficient and durable decision rules: A reformulation" Econometrica, 53, 817-835.

[6] Dequiedt, V., 2006. "Ratification and veto constraints in mechanism design" working paper.

[7] Forges, F., 1999. "Ex post individually rational trading mechanisms" in Current Trends in Economics, A. Alkan, C. Aliprantis, and N. Yannelis (eds.), Springer-Verlag, 157-175.

[8] Goemans, H. E. and M. Fey, 2009. "Risky but rational: War as an institutionally induced gamble," Journal of Politics, 71, 35-54.

[9] Holmstrom, B. and R. Myerson, 1983. "Efficient and durable decision rules with incomplete information" Econometrica, 51, 1799-1819.

[10] Jehiel, P., B. Moldovanu, and E. Stacchetti, 1996. "How (not) to sell nuclear weapons" American Economic Review, 86, 814-829.

[11] Jehiel, P., B. Moldovanu, and E. Stacchetti, 1999. "Multidimensional mechanism design for auctions with externalities" Journal of Economic Theory, 85, 258-293.

[12] Kamenica, E. and M. Gentzkow, forthcoming. "Bayesian persuasion" American Economic Review.

[13] Kim, J., 2008. "The value of an informed bidder in common value auctions" Journal of Economic Theory, 143, 585-595.

[14] Laffont, J. J. and D. Martimort, 1997. "Collusion under asymmetric information" Econometrica, 65, 875-911.

[15] Laffont, J. J. and D. Martimort, 2000. "Mechanism design with collusion and correlation" Econometrica, 68, 309-342.

[16] Lagunoff, R. D., 1995. "Resilient allocation rules for bilateral trade" Journal of Economic Theory, 66, 463-487.

[17] Maskin, E. and J. Tirole, 1990. "The principal - agent relationship with an informed principal: The case of private values" Econometrica, 58, 379-409.

[18] Maskin, E. and J. Tirole, 1992. “The principal - agent relationship with an informed principal II: Common values” Econometrica, 60, 1-42.

[19] Matthews, S. and A. Postlewaite, 1989. “Preplay communication in two-person sealed-bid double auction” Journal of Economic Theory, 48, 238-263.

[20] Motta, A. 2010. “Collusion and selective supervision” working paper.

[21] Myerson, R.B., 1983. Mechanism design by an informed principal. Econometrica 51, 1767–1797.

[22] Myerson, R. B., 1991. “Game theory: Analysis of conflict” Harvard University Press: Cambridge MA, London England.

[23] Palfrey, T. and S. Srivastava, 1991. “Efficient trading mechanisms with pre-play communication” Journal of Economic Theory, 55, 17-40.

[24] Philippon, T. and V. Skreta, 2010. “Optimal interventions in markets with adverse selection” working paper.

[25] Tan, G. and O. Yilankaya, 2007. “Ratifiability of efficient collusive mechanism in second-price auctions with participation costs” Games and Economic Behavior, 59, 383-396.

[26] Tirole, J, forthcoming. “Overcoming adverse selection: How public intervention can restore market functioning” American Economic Review.