RECIPROCAL RELATIONSHIPS AND MECHANISM DESIGN

GORKEM CELIK AND MICHAEL PETERS

Abstract. We study an incomplete information game in which players can coordinate their actions by contracting among themselves. We model this relationship as a *reciprocal contracting* procedure where each player has the ability to make commitments contingent on the other players' commitments. We differ from the rest of the literature on reciprocal contracting by assuming that punishments can't be enforced in the event that cooperation breaks down. We fully characterize the outcomes that can be supported as *perfect* Bayesian equilibrium outcomes in such an environment. We use our characterization to show that the set of supportable outcomes with reciprocal contracting is larger than the set of outcomes available in a centralized mechanism design environment in which the mechanism designer is constrained by his inability to enforce punishments against non-participants. The difference stems from the players' ability in our contracting game to convey partial information about their types at the time they offer contracts. We discuss the implications of our analysis for modeling collusion between multiple agents interacting with the same principal.

KEYWORDS: Conditional contracts; Default game; Signaling; Collusion.

1. Introduction

Consider a couple contemplating marriage. This couple regards marriage as a mutual commitment which would constrain the choices each of them can make in life. Before finalizing their decision to get married, they have to discuss many parameters of their new life together including where they will get married, where they will live, how they will finance their future, if they will have children, etc. There is a possibility that they cannot agree on some of these parameters and therefore they do not get married, at least for the time being. In this case, there is no commitment made and the potential spouses have access to the same set of choices as they had before they started talking about marriage. However, once the question of marriage arises, it would be naïve to imagine that its alternative is the continuation of the status quo for this couple. For instance, if one party held a less conciliatory position

Date: June 26, 2014.
Acknowledgments to be added.
Celik: ESSEC Business School and THEMA Research Center, celik@essec.fr ;
Peters: University of British Columbia, peters@econ.ubc.ca.

during the failed marriage deliberations, the other one may start questioning the reliability of his/her partner. The statements made by these two people are not only the potential building blocks of a marriage agreement between them but also the means of *signaling* who they are to their partner. The information revealed by these signals is relevant in shaping the outside option for the marriage as it is relevant in devising the marriage.

In the language of this paper, the two partners above are the players of an incomplete information *default game*, and marriage is a *contract* that they can sign with the purpose of constraining / coordinating their actions in this game. Other examples captured by this setting include competing firms deliberating a cartel agreement, bidders discussing the formation of an auction bidding ring, and disputing governments negotiating a peace settlement. Our objective is to find out what outcomes the players can achieve in the default game with the help of contracts.

In what follows, we endow each player with the ability to communicate with other players and to make commitments based on these communications. In particular, we let each player write a *reciprocal contract* which conditions his default game actions directly on the contracts of the other players. The exact nature of how this conditioning works is explained in detail below. In a nutshell, if all the reciprocal contracts agree with one another, then they implement some kind of a cooperative action. If they do not agree, then the contracts are void and each player is free to choose any action he wants in the default game.

Our paper is connected to the recent literature on contracts which can condition on one another. This idea is given precise content by Tennenholtz (2004) who formalizes a game played by computer programs, each of which conditions its action on some other program. Kalai et al. (2010) uses the same idea to describe a two player contracting game of complete information where the Nash equilibrium set coincides with the set of joint mixtures over actions for which each player receives at least his minmax payoff.[1] Forges (2013) and Peters (2013) extend the characterization of conditional contracting outcomes to incomplete information games under the solution concepts of Bayesian equilibrium and perfect Bayesian equilibrium respectively. To avoid infinite regress in the conditioning of the contracts, these papers restrict the sets of contracts available to the players. Peters and Szentes (2012) follow a different approach and consider a *universal set* of conditional contracts which can be written in finite text.

What differentiates our paper from this earlier literature is that the players in our reciprocal contracting game do not have enough commitment power to enforce *punishments* on the players deviating from equilibrium behavior. Essentially players

[1]Bachi, Ghosh, and Neeman (2013) introduce the possibility of breaking the commitment made in the conditional contract by incurring a *deception* cost.

can bind their actions when all players unanimously agree to some course of action. Otherwise, any kind of disagreement leads to a complete breakdown with every player choosing his action in a sequentially rational way during the default game. For instance, if a firm does not want to participate in a cartel agreement, we do not allow some other firm to cut its price to zero just as a punishment for the non-participating firm. Instead, we require that each firm chooses the default game action that maximizes its expected payoff in case of a disagreement.[2] The limitation on the punishments introduces a new function for the conditional contracts which has been absent in the earlier literature: in our setting, each player's contract can be used as a signaling device in order to manipulate the disagreement payoffs of the players.

Our main result is the characterization of all the outcome functions that can be supported as (perfect Bayesian) equilibrium outcomes in our reciprocal contracting game. An important subset of these outcomes is supported by equilibria which do not involve any information revelation by the players during the negotiation of the contract. We argue that these *pooling* equilibrium outcomes coincide with the outcomes that can be sustained by a *mechanism designer*, who is *constrained* to offer a contract that is acceptable to all players regardless of their private information and who cannot influence the play in the default game when a player *unexpectedly* rejects the contract.

We also show that the reciprocal contracting game has *separating* and partially separating equilibria which support outcomes that this constrained mechanism designer cannot. The ability to reveal partial information during the contracting process changes the outside option of players in the default game, should they decide not to cooperate. In a separating equilibrium, a player could still trigger the non-cooperative play of the default game by not reciprocating with the other players. In this case, the default game would be played under the updated beliefs on the types of the non-deviating players because of the signaling that occurs through the contract offers that these players make. As a consequence, the payoff that the deviating player receives is the expectation of his *non-cooperative payoff* against the various posterior beliefs he might face. This expectation could well be lower than what this player would have received in the default game played under the prior beliefs.

Our paper also relates to mechanism design papers by Cramton and Palfrey (1995), Caillaud and Jehiel (1998), Tan and Yilankaya (2007), Jullien, Pouyet, and Sand-Zantman (2011). In these papers, the design problem is a centralized one where an uninformed constrained designer offers a contract to the players. As in our paper, a player can trigger the non-cooperative play of the default game by refusing this

[2]See also Koessler and Lambert-Mogiliansky (2013) for an example to a conditional commitment (to transparency in order to fight corruption) technology which does not involve non-credible punishments for the deviators.

contract. A player's acceptance decision is the only means of signaling his type to the others before the contract takes effect. This strand of the literature studies equilibria where all types of all players accept the designer's offer. Therefore belief updates are possible only off the equilibrium path, when a player refuses the designer's contract unexpectedly. By contrast, in our decentralized design setting, we need to account for on the equilibrium path belief updates, which are supported by the type dependence of the players' contracts.[3] The information revealed at the contracting stage affects players' incentives on the equilibrium path as well as off of it. For this reason, our characterization result will refer to incentive compatibility constraints which are different than the standard interim constraints.

In earlier work (Celik and Peters, 2011), we show that equilibrium path signaling can be sustained even in a constrained centralized design setting. This is possible if players' acceptance decisions of the designer's contract depend on their types.[4] In this earlier work, we demonstrate that the equilibrium path belief update opportunities enlarge the set of available outcomes. However, the additional outcomes achieved through signaling require the contract to be rejected with a strictly positive probability. In the current setting, players can signal their type through their contract offers – in a manner similar to Peters and Szentes (2011) – without having to reject agreement with positive probability.[5] This expands the set of supportable outcomes beyond those described in Celik and Peters (2011).

The rest of the paper is organized as follows. In Sections 2 and 3, we introduce the default game and the reciprocal contracting game. In Section 4, we describe the incentive constraints and characterize the equilibrium outcomes of the reciprocal contracting game by referring to these constraints. In Section 5, we discuss the implications of our analysis for modeling collusion between multiple agents interacting with the same principal. We show that, if the agents are colluding through reciprocal contracts, the principal can implement outcomes which are deemed to be prone to collusion by the earlier literature. In Section 6, we consider a private values environment under the single crossing and transferable payoff conditions. For this frequently studied setting, we show that the incentive constraints in our characterization result can be simplified further. Moreover the discrepancy between our results on collusion and those of the earlier literature disappears. Section 7 is the conclusion. We relegate the proof of the characterization theorem to the Appendix.

[3]See Maskin and Tirole (1990 and 1992) for another example to the signaling potential of contracts offered by informed parties.

[4]In a similar vein, Philippon and Skreta (2012) and Tirole (2012) argue that a failing firm's refusal to participate in a government sponsored bailout plan may signal its confidence in the performance of its assets.

[5]Unlike Peters and Szentes (2011), we allow action spaces to be continuous and fully incorporate mixed strategies and stochastic contracts.

2. The Default Game

In an environment with incomplete information, I is the set of players. We refer to the private information of a player as his type. Each player $i \in I$ has a finite type set T_i. The actions available to player i in this game are elements of set A_i, which is a closed subset of an Euclidean space with finite dimension. In standard notation A and T are cross product spaces representing all players' actions and types respectively. Notice that, by this construction, A is a subset of an Euclidean space $\mathbb{R}^K$.[6]

In order to retain the idea that sensible off the equilibrium path beliefs about a player can only change after that player himself has deviated, we assume from the outset that types are independently distributed. Player i's type is distributed with respect to the prior distribution β_i^0. We define $\beta^0 = \{\beta_i^0\}_{i \in I}$ as the collection of these priors. $\beta_i^0(t_i)$ is the probability that player i has type $t_i \in T_i$ under the prior distribution. Similarly, $\beta^0(t) = \prod_{i \in I} \beta_i^0(t_i)$ denotes the probability that the realization of the type profile is $t = \{t_i\}_{i \in I} \in T$.

Preferences of player i are given by the payoff function $u_i : A \times T \to \mathbb{R}$. Players have expected utility preferences over lotteries. If $q \in \Delta A$ is a randomization over action profiles and $t \in T$ is a type profile, then $u_i(q, t)$ refers to the associated expected utility with a slight abuse of notation. An **outcome function** is a mapping from type profiles into randomizations over action profiles $\omega : T \to \Delta A$.

In the absence of a technology to write down and commit to mechanisms, the set of players, the action sets, the type sets, and the payoff functions define a Bayesian game together with the prior distribution $\beta^0 = \{\beta_i^0\}_{i \in I}$. The fact that players may choose different actions under different beliefs is central to our analysis. Therefore we study this game under an arbitrary distribution $\beta = \{\beta_i\}_{i \in I}$, rather than the prior distribution. As in the definition of the prior distribution, β_i is an element of ΔT_i and $\beta_i(t_i)$ is the probability that player i has type t_i under this distribution.

We refer to the collection $I, \{T_i\}_{i \in I}, \{A_i\}_{i \in I}, \{u_i\}_{i \in I}$, and β as the default game under belief β. When playing this game, each type of each player chooses his action to maximize his expected payoff. Accordingly, collection of functions $\{q_i(\cdot|\beta)\}_{i \in I}$ constitutes a **Bayesian equilibrium** of the default game under belief β if each action in the support of randomization $q_i(t_i|\beta)$ is a solution to

$$\max_{a_i \in A_i} \mathbb{E}_{t_{-i}|\beta_{-i}} \left[u_i \left(a_i, q_{-i}(t_{-i}|\beta), t_i, t_{-i} \right) \right],$$

[6] If k_i is the dimension of action set A_i for each player i, then K is smaller than or equal to $\sum_{i \in I} k_i$. Our construction allows for the possibility that A_i is finite for each i. In this case, A can be considered as a finite subset of real line $\mathbb{R}$.

for each type $t_i \in T_i$ of each player $i \in I$. The operator $\mathbb{E}_{t_{-i}|\beta_{-i}}$ stands for the expectation over the values of t_{-i} given belief β_{-i}.[7]

We restrict attention to default games for which a Bayesian equilibrium exists. Existence of equilibrium is immediate for games with finite action sets. For simplicity of exposition, we assume further that there is a **unique** Bayesian equilibrium of the default game. We can extend the analysis to games with multiple equilibria with a slightly more complicated statement of the incentive constraints below. Alternatively, the unique equilibrium, to which we refer, can be thought as the equilibrium chosen (among possibly multiple equilibria) by some selection criteria.[8]

Suppose that $\{q_i(\cdot|\beta)\}_{i\in I}$ is the Bayesian equilibrium of the default game under belief β. We define the **non-cooperative payoff** V_i as the function that maps the types of player i and the beliefs into expected equilibrium payoffs:

$$V_i(t_i, \beta) = \mathbb{E}_{t_{-i}|\beta_{-i}}\left[u_i\left(q_i(t_i|\beta), q_{-i}(t_{-i}|\beta), t_i, t_{-i}\right)\right]. \tag{2.1}$$

2.1. **Example: The Cournot Default Game.** Consider a game played by two *quantity setting* firms (players) who have the technology to produce the same homogenous good. Each player has a constant unit production cost which is his private information. Unit cost (type) of player 1 is either 48 or 56. Unit cost of player 2 is either 52 or 80. These types are independently distributed for the players. The inverse demand function for the good they produce is given as $P = 80 - (y_1 + y_2)$, where P is the price and y_1, y_2 are the production levels of players 1 and 2. Assuming that each player is an expected profit maximizer, we can write player i's utility function as $u_i(y_i, y_j, t_i) = [80 - (y_i + y_j) - t_i]\, y_i$. Since each player has a binary type set, we can represent a probability distribution over the types of a player with a single probability. Let β_i denote the probability that player i has a lower cost type. That is, β_1 is the probability that player 1 has type (48) and β_2 is the probability that player 2 has type (52). Under any pair of beliefs (β_1, β_2), this game has a unique Bayesian equilibrium. The resulting equilibrium production and expected payoff levels are reported below:

$$\begin{array}{ll} y_1(48|\beta_1,\beta_2) = \frac{16-8\beta_2+\beta_1\beta_2}{1-0.25\beta_2} & V_1(48,\beta_1,\beta_2) = \left(\frac{16-8\beta_2+\beta_1\beta_2}{1-0.25\beta_2}\right)^2 \\ y_1(56|\beta_1,\beta_2) = \frac{12-7\beta_2+\beta_1\beta_2}{1-0.25\beta_2} & V_1(56,\beta_1,\beta_2) = \left(\frac{12-7\beta_2+\beta_1\beta_2}{1-0.25\beta_2}\right)^2 \\ y_2(52|\beta_1,\beta_2) = \frac{8-2\beta_1}{1-0.25\beta_2} & V_2(52,\beta_1,\beta_2) = \left(\frac{8-2\beta_1}{1-0.25\beta_2}\right)^2 \\ y_2(80|\beta_1,\beta_2) = 0 & V_2(80,\beta_1,\beta_2) = 0 \end{array} \tag{2.2}$$

[7]In standard notation, subscript $-i$ refers to the collection of one variable for each player other than player i. For instance, $t_{-i} = \{t_j\}_{j\in I-\{i\}}$.

[8]A similar *uniqueness* assumption appears in the recent work of Hagenbach, Koessler, and Perez-Richet (2014) with the same justification.

Regardless of the beliefs, the high cost type of player 2 produces zero output, since his production cost is higher than the price. For the other three types (types (48) and (56) of player 1 and type (52) of player 2), the equilibrium behavior and the expected payoff depend on the beliefs under which the game is played. Notice that player i's Bayesian equilibrium output and expected payoff are weakly decreasing in his belief that his rival (player j) has the lower cost (β_j), and weakly increasing in player j's belief that player i has the lower cost (β_i).

We now introduce an extension of this game by allowing monetary transfers between the players. We assume that in addition to setting his production level y_i, each player can also make a non-negative transfer z_i to the other player. Players maximize their utility net of the transfers.[9] These monetary transfers will be useful instruments for agreements between the players. However, in the absence of a binding mechanism, the equilibrium behavior for each player is making a zero transfer. Therefore the equilibrium payoff functions we gave above are the non-cooperative payoffs for the extended game as well. This game, which we call the **Cournot default game**, is a modified version of the example studied by Celik and Peters (2011). In what follows, we will refer to the Cournot default game several times in order to illustrate some key points of our analysis.

3. The Reciprocal Contracting Game

We model the contracting process as a slightly modified version of the reciprocal contracting game introduced by Peters (2013). Reciprocal contracts are a convenient way to model the kind of contracting situations we have in mind because, as shown by Peters, their equilibrium outcomes can mimic the equilibrium outcomes of a broad variety of different contracting games. This is important for problems like marriage, cartel formation, or collusion where it is difficult to know exactly how contracts are being negotiated. Reciprocal contracting provides a way to understand the entire spectrum of behavior that is supportable with contracts. Unlike Peters, we assume that the players cannot commit to a course of action unless the players unanimously agree on it.

The contracting process takes place in two rounds. In the first round, players offer *contracts*. These contracts determine a *mechanism* for each player, committing this player to an action contingent on messages that will be sent in the second round. The key feature of this process is the dependence of a player's mechanism on the mechanisms of the other players. This conditioning can either be explicit, as in Peters and Szentes (2012), or implicit as in the contracting game we explain below.

[9]Player i's utility net of the transfers is $\tilde{u}_i(y_i, y_j, z_i, z_j, t_i) = [80 - (y_i + y_j) - t_i]\, y_i - z_i + z_j$, where z_i and z_j are the monetary transfers made by firms i and j respectively.

In line with the literature, we define a mechanism for a player as a mapping from the cross product of the message sets into the actions that this player can take. The contracting game relies on the class of **direct mechanisms**. A player's message to a direct mechanism consists of a *type report* t_i and a *correlating message* n_i which is a K dimensional vector whose components are real numbers on the interval $[0, 1]$. Recall that number K represents the dimension of the set of action profiles A. A direct mechanism for player i transforms the type reports and the correlating messages of the $|I|$ players into a default game action that this player will take. Let N be the set of $|I| \times K$ matrices consisting of real numbers on the interval $[0, 1]$. Then a direct mechanism for player i is formally defined as:

$$m_i : T \times N \to A_i.$$

Notice that direct mechanisms are defined as deterministic mechanisms, i.e., each message profile is mapped into a single action instead of a randomization over actions. When proving our characterization theorem, we will explain how the correlating message vectors would generate a *jointly controlled lottery*, which supports randomizations over actions. Moreover we will show that these randomizations may be correlated across players as well.[10] M_i denotes the set of all direct mechanisms for player i.[11]

In the first round of the contracting game, each player offers a **reciprocal contract**. A reciprocal contract gives a player the opportunity to make a revelation about his type and the possibility of committing to a mechanism contingent on the revelations made by the others. Recall that types of players are independently distributed. Before the first round of the contracting procedure, all the other players believe that prior β_i^0 governs the distribution of player i's type. After observing the revelation made by player i, the prior belief on this player is updated to a posterior belief. We need the set of possible revelations by each player to be large enough to support any possible posterior distribution. To make it explicit that players signal their types with their contract offers, we model each player's revelation as announcing a distribution of his types.[12] Formally, a reciprocal contract for player i consists of a

[10]See Aumann and Maschler (1995) for jointly controlled lotteries, used in the context of repeated games. In our work, correlating message vectors stand as a proxy for a *public randomization device*. For settings in which a public randomization device exists, we could define a direct mechanism for player i as a mapping from the type reports and the realizations of the commonly observed random variable into the actions of player i.

[11]In what follows, we assume that the players use these direct mechanisms as the building blocks of their contracts. Our analysis extends to the case where the mechanisms available to each player constitute a superset of the direct mechanisms.

[12]Our analysis extends to the case where the revelations available to each player constitute a superset of the set of distributions of his types.

revelation $\hat{\beta}_i \in \Delta T_i$ and a list of potential direct mechanisms δ which is represented by a mapping from revelations of all players into profiles of direct mechanisms

$$\delta : \times_{j \in I} \Delta T_j \rightarrow \times_{j \in I} M_j.$$

These contracts determine the players' mechanisms as follows: if all contracts include the same list δ, then the mechanisms are indeed pinned down by how this function maps the submitted revelations of players into a profile of mechanisms. That is, if the players' offers in the first round agree on function δ, than the direct mechanism $\delta_i\left(\hat{\beta}\right)$ determines the mechanism that player i will follow in the second round, where δ_i is the i^{th} component of function δ and $\hat{\beta} = \left\{\hat{\beta}_i\right\}_{i \in I}$ is the profile of revelations made in the first round. However, if there is at least one player who offered a contract containing a different list δ than did the other players, then no mechanism takes effect. Instead, each player i chooses his default game action non-cooperatively. Reciprocal contracts are intended to look like mutual agreements – if all players agree, cooperation occurs. Otherwise, when a player does not reciprocate, as a "punishment" to this player, the default game is played non-cooperatively.

A contract offer, by construction, leads to a specific commitment for a player. Yet this offer does not necessarily resolve all of the player's uncertainty. He does not know what he himself has committed to until he sees all of the other contracts. If he expects the other players to offer contracts that list the same array of mechanisms δ that he does, then he believes that the first round revelations of all players will determine his commitment as well as the commitments of the others.

The reciprocal contracting process induces an imperfect information game with two stages. We base our analysis of this sequential game on the solution concept of perfect Bayesian equilibrium, which consists of strategies and beliefs satisfying the conditions below:

i) In round 1, each type of each player i chooses his contract to maximize his expected continuation payoff.

ii) After observing player i's contract offer, other players update their beliefs on his type. On the equilibrium path, the belief updates are governed by the Bayes rule. Off the equilibrium path, all players other than player i share a common posterior on player i's type.[13]

iii) In round 2, on the equilibrium path or on the continuation games reachable by unilateral deviations from the equilibrium play,[14] each type of each player i chooses

[13]In other words, after observing player i's off the equilibrium path behavior, all the other players update their beliefs in the same way. This assumption is consistent with Fudenberg and Tirole's (1991) definition of perfect Bayesian equilibrium.

[14]Our definition of perfect Bayesian equilibrium demands sequential rationality of strategies only for continuation games which are either on the equilibrium path or accessible by unilateral

his message to the mechanisms (if all contract offers include the same list of direct mechanisms) or his default game action (if some player offers a different list) in order to maximize his expected continuation payoff, given the updated beliefs.

4. Incentive Constraints

The main objective in this paper is providing a characterization of the outcome functions which are supportable as the perfect Bayesian equilibrium outcomes of the reciprocal contracting game. With our first result, we show that it is sufficient to restrict attention to a specific class of equilibria for this characterization.

Proposition 1. *If ω is a perfect Bayesian equilibrium outcome function of the reciprocal contracting game then it is also supportable as the outcome of a perfect Bayesian equilibrium of this game where*
i) players reciprocate*: all types of all players submit a unique list of mechanisms δ^* as part of their contracts in round 1;*
ii) revelations are accurate*: on the equilibrium path, after observing revelation $\hat{\beta}_i \in \Delta T_i$ by player i, all the other players update their posterior belief to $\hat{\beta}_i$;*
iii) type reports are truthful*: on the equilibrium path, all players report their types truthfully to the mechanisms in round 2;*
iv) correlating message vectors are uniformly distributed*: on the equilibrium path, for each player, each component of the correlating message vector is uniformly distributed on the interval $[0, 1]$ regardless of the player's type and the posterior beliefs.*

We provide the proof of this proposition in the Appendix together with the proof of our characterization theorem.

Property (i) above follows from a familiar argument. Suppose there exists an equilibrium where some types of some players do not reciprocate, i.e., they submit a list of mechanisms other than δ^*. The very same equilibrium outcome could have been supported by an alternative equilibrium where all types of all players agree on an "extended" list of mechanisms. This extended list replicates the non-cooperative play of the default game following the non-reciprocating behavior in the original equilibrium. Property (iii) is a direct implication of the revelation principle. Property (iv) points to the fact that correlating messages are used in order to generate jointly controlled lotteries (to be used as public randomization devices) in this class of equilibria.

---

deviations from the equilibrium behavior. As Peters and Troncoso Valverde (2013) demonstrate, optimality of strategies in **all** nodes of the extensive form game is not possible to achieve: there may be continuation games triggered by players agreeing on direct mechanisms which do not have an equilibrium in pure or mixed strategies.

The intuition for property (ii) follows from the revelation principle as well. If all types of all players submit the same list of mechanisms δ^*, their revelation messages are the only means of separating different types of players on the equilibrium path. Each revelation by player i will lead to a potentially different posterior on his types. In the class of equilibria defined by the proposition above, the equilibrium path revelations are re-labeled in such a way that they match the posterior beliefs they generate.

In an equilibrium which satisfies the properties above, a player can deviate from equilibrium play either by refusing to reciprocate with the other players, or by making an inaccurate revelation about his type in round 1, or by misreporting his type in round 2. The outcome functions must satisfy certain incentive constraints for these deviations not to be profitable. We describe these constraints below and discuss how they relate to the more familiar versions invoked in the earlier literature. Then we provide a formal characterization of equilibrium outcome functions by referring to the described constraints.

4.1. **Individual Rationality.** In an equilibrium where all players are expected to reciprocate, any player can trigger the non-cooperative play of the default game by offering a different list of mechanisms. As a result of this unilateral deviation, each player i receives the non-cooperative payoff $V_i\left(t_i, \beta_i, \beta_{-i}\right)$ defined in (2.1). For this deviation not to be profitable, each type of each player must expect an equilibrium payoff weakly higher than his non-cooperative payoff. This consideration yields the individual rationality constraints.

After a player's refusal to reciprocate, the beliefs on the players' types need not remain the same as the prior beliefs. First of all, as a result of the refusal of player i, the other players may update their belief regarding the type of this player from prior β_i^0 to some posterior β_i^{no}. We refer to the collection of these beliefs $\beta^{no} = \{\beta_i^{no}\}_{i\in I}$ as the **refusal beliefs.** In the construction of an equilibrium where all players reciprocate, refusal beliefs are arbitrary. This is due to the fact that standard solution concepts such as perfect Bayesian equilibrium do not put much restriction on beliefs off the equilibrium path.[15]

In addition to changing their beliefs on the type of a deviating player, the participants of the reciprocal contracting game may update their beliefs on the non-deviating players as well. Recall that players are allowed to make revelations about

[15] It is possible to suggest a refinement of perfect Bayesian equilibrium by imposing additional requirements on these refusal beliefs. For instance, setting $\beta_i^{no} = \beta_i^0$ for all i amounts to assuming *passive beliefs* (Cramton and Palfrey, 1990). Alternatively, one may assume that the support of refusal beliefs consists only of the types that are not strictly worse off by rejecting to reciprocate. This refinement leads to the concept of *ratifiability* (Cramton and Palfrey, 1995). We will continue our analysis without imposing such a refinement and allowing for arbitrary rejection beliefs.

their private information as part of their contract offers. As a result, their equilibrium play may indeed unveil information on their types. Any player who contemplates deviating should understand the impact that other players' revelations will have.

Consider type t_i of player i's decision to reciprocate with the others in the first round of the game. This player knows that the others will update their beliefs to β_i^{no} if he does not reciprocate. He also comprehends that after observing the contract offers, the beliefs on the other players' types will be updated to some posterior β_{-i}. Recall that the non-cooperative payoff $V_i\left(t_i, \beta_i^{no}, \beta_{-i}\right)$ yields the continuation payoff of player i from the non-cooperative play of the default game under these beliefs. There is one more complication in the analysis of player i's decision to reciprocate. Player i has to make this decision before he learns the other contracts and observes the revelations by the other players. Therefore, at the time he makes the decision, player i does not know the exact realization of the posterior β_{-i}. However, the equilibrium strategies of the other players reveal the distribution over the possible posteriors.

We represent a distribution over the posteriors on the types of player i with notation $\Pi_i \in \Delta\left(\Delta T_i\right)$. Suppose this distribution is indeed generated by revelations made by player i on the equilibrium path. In this case, the Bayes rule implies that the expectation over the posteriors equals the prior distribution: $\mathbb{E}_{\beta_i|\Pi_i}\beta_i = \beta_i^0$. Following Kamenica and Gentzkow (2011), we call distribution Π_i **Bayes plausible** when it satisfies this property**.** If Π_i is Bayes plausible for each player i, then we refer to the collection $\Pi = \{\Pi_i\}_{i\in I}$ as a **posterior system.**

Suppose the outcome function ω is supportable by an equilibrium where all players reciprocate with each other by submitting the same list of mechanisms. Then each player must have the incentive not to deviate by refusing to reciprocate. This is ensured with the following individual rationality condition. Under the refusal beliefs β^{no} and the posterior system Π, outcome function ω is **individually rational** if

$$\mathbb{E}_{t_{-i}|\beta_{-i}^0}\left\{u_i\left(\omega\left(t\right), t\right)\right\} \geq \mathbb{E}_{\beta_{-i}|\Pi_{-i}}\left\{V_i\left(t_i, \beta_i^{no}, \beta_{-i}\right)\right\} \tag{4.1}$$

for all t_i and all i.

Consider an equilibrium of the contracting game, where no relevant information is revealed with the equilibrium contract offers. We can represent the resulting information structure with a posterior system Π which puts unit mass on the prior distribution β^0. Under this system, the right hand side of (4.1) boils down to $V_i\left(t_i, \beta_i^{no}, \beta_{-i}^0\right)$:

$$\mathbb{E}_{t_{-i}|\beta_{-i}^0}\left\{u_i\left(\omega\left(t\right), t\right)\right\} \geq V_i\left(t_i, \beta_i^{no}, \beta_{-i}^0\right) \tag{4.2}$$

for all t_i and all i.

For player i, payoff $V_i\left(t_i, \beta_i^{no}, \beta_{-i}^0\right)$ corresponds to the non-cooperative play of the default game under no additional information other than the prior beliefs. Other posterior systems would have given player i more information on the types of the other players. Nevertheless, $V_i\left(t_i, \beta_i^{no}, \beta_{-i}^0\right)$ does not necessarily constitute a lower bound on the right hand side of (4.1). In other words, the reservation utility of player i may decrease with the level of information revealed by the contracts offered in the first round. We will demonstrate this point with the help of the Cournot default game introduced earlier.

4.1.1. ***Individual Rationality and the Cournot Default Game.*** Suppose that the two players of the Cournot default game are negotiating over a cartel agreement by using the reciprocal contracting process we described. If these players can agree on the cartel, their agreement will determine their type dependent production levels and the monetary transfers they will make to each other. Each player has the option to refuse to participate in the cartel. By doing so, the player triggers the non-cooperative play of the Cournot default game. Before they start the negotiations, each player's belief on the type of the other player is uniform. That is, the prior beliefs are given as $\beta_1^0 = \beta_2^0 = 0.5$.

What payoff would a player expect from not participating in the cartel? The non-cooperative payoff functions we report in (2.2) indicate that Player 2 with cost (80) would receive zero payoff from the default game regardless of the beliefs. The non-cooperative payoffs of the other types of players decrease in the likelihood that they are perceived to have a higher cost. This means that a larger set of outcome functions will be classified as individually rational if a non-participating player is believed to be the highest cost type with probability one. This is ensured by the following refusal beliefs: $\beta_1^{no} = \beta_2^{no} = 0$.

What about the beliefs on the type of the non-deviating player? Since these beliefs are equilibrium path beliefs, they should be equal to the prior belief in expectation. That is, if Π_i is the distribution over the equilibrium beliefs on the types of player i, then $\mathbb{E}_{\beta_i|\Pi_i}\beta_i = \beta_i^0 = 0.5$. As long as it satisfies this Bayes plausibility condition, any Π_i is supportable as a distribution over the equilibrium beliefs.

Let us start with considering the non-cooperative payoff function of player 2 with cost (52). This type's non-cooperative payoff $V_2(52, \beta_1, \beta_2)$ is convex in β_1, implying that his expected default game payoff would increase in the information he receives on player 1's type. Therefore the right hand side of (4.1) would be minimized if the distribution of posteriors assigns unit mass to the prior distribution $\beta_1^0 = 0.5$. We label this degenerate distribution of posteriors as Π_1^*. Under the refusal belief $\beta_2^{no} = 0$ and the prior belief $\beta_1^0 = 0.5$, the non-cooperative payoff for type (52) of player 2 is $V_2(52, 0.5, 0) = 7^2 = 49$. Player 2 with type (52) will not accept the cartel agreement if he receives a payoff lower than this figure.

Now we turn our attention to player 1. This player's non-cooperative payoff is not convex in β_2 for either one of his two types. This non-convexity indicates that this player's expected payoff can be reduced by revealing information to him on player 2's type. To see this, first suppose that the distribution over the posteriors of types of player 2 assigns unit mass to the prior. Under the refusal belief $\beta_1^{no} = 0$ and the prior belief $\beta_2^0 = 0.5$, the non-cooperative payoff of player 1 is $V_1(48, 0, 0.5) = \left(\frac{16-4}{1-0.125}\right)^2 \cong 188.08$ for type (48) and $V_1(56, 0, 0.5) = \left(\frac{12-3.5}{1-0.125}\right)^2 \cong 94.37$ for type (56). If player 1 believes that he will not receive any additional information about his rival, these numbers determine his type dependent reservation utility.

As mentioned above, thanks to the non-convexity of function V_1, one could reduce this reservation utility by revealing player 1 some information about the type of player 2. For instance, consider the distribution of posteriors Π_2^* which assigns probability 3/8 to posterior $\beta_2 = 0$ and probability 5/8 to posterior $\beta_2 = 0.8$. Notice that Π_2^* is Bayes plausible since $\mathbb{E}_{\beta_2|\Pi_2^*}\beta_2 = 0.5$. In order to support Π_2^* as the distribution of posteriors in the reciprocal contraction game, it would suffice to construct an equilibrium where type (52) of player 2 makes the revelation $\hat{\beta}_2 = 0.8$ with probability one and type (80) of player 2 randomizes between $\hat{\beta}_2 = 0.8$ and $\hat{\beta}_2 = 0$ with probabilities 1/4 and 3/4 respectively. Under Π_2^*, the expected value of the non-cooperative payoff for the two types of player 1 are as below:

$$\begin{aligned}
\mathbb{E}_{\beta_2|\Pi_2^*}V_1(48, 0, \beta_2) &= \frac{3}{8}V_1(48, 0, 0) + \frac{5}{8}V_1(48, 0, 0.8) = \frac{3}{8}256 + \frac{5}{8}144 = 186, \\
\mathbb{E}_{\beta_2|\Pi_2^*}V_1(56, 0, \beta_2) &= \frac{3}{8}V_1(56, 0, 0) + \frac{5}{8}V_1(56, 0, 0.8) = \frac{3}{8}144 + \frac{5}{8}64 = 94.
\end{aligned}$$

Notice that both numbers are lower than the non-cooperative payoffs corresponding to the alternative scenario where the default Cournot game is played under the prior belief $\beta_2^0 = 0.5$. Figure 1 illustrates this for type (56) of player 1.[16] In fact, it follows the analysis in Celik and Peters (2011) that any distribution other than Π_2^* would result in a strictly higher expected non-cooperative payoff for at least one of the types of player 1.[17]

So far, we defined two distributions over the posteriors on the types of player 1 (Π_1^*) and player 2 (Π_2^*). These two constitute a posterior system $\Pi^* = \{\Pi_1^*, \Pi_2^*\}$. Under the posterior system Π^* and refusal beliefs $\beta_i^{no} = 0$, player 1 expects to receive payoff 186 for type (48) and payoff 94 for type (56) from the non-cooperative play of the Cournot default game. Similarly, player 2's non-cooperative payoff is 49 for type (52) and 0 for type (80). These figures pin down the reservation utility levels on the right hand sides of constraints in (4.1).

[16]For ease of demonstration, figure is not drawn to scale.

[17]Distribution Π_2^* minimizes the expected value of the non-cooperative payoff of type (56) of player 1, since $\mathbb{E}_{\beta_2|\Pi_2^*}V_1(56, 0, \beta_2) = 94$ is the value of the *biconjugate* of function $V_1(56, 0, \beta_2)$ at $\beta_2 = 0.5$. Any other distribution of posteriors would support a strictly higher non-cooperative payoff than 94 for type (56).

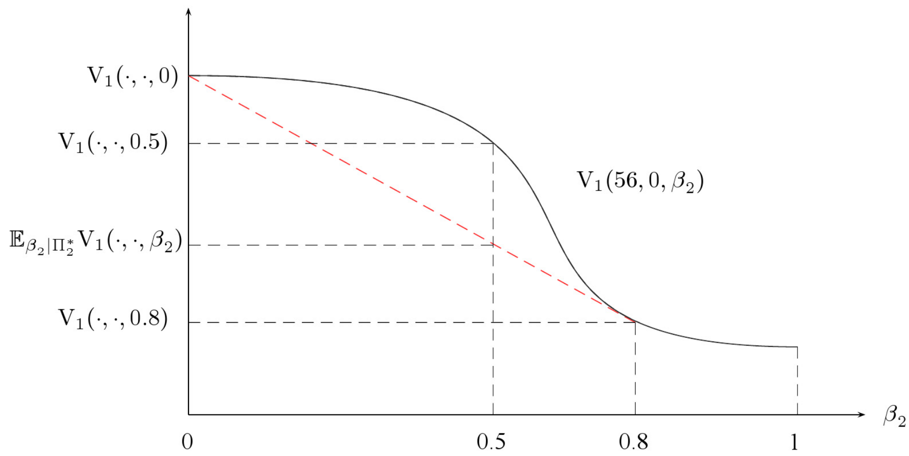

FIGURE 1. Non-convexity of the non-cooperative payoff function

The interesting point about the above construction is the decline in player 1's expected non-cooperative payoff as he gets superior information on the type of the other player through the latter's revelations. What is the reason for this seemingly negative value for information? The answer lies in the observation that, in this setting, it is not possible to single out one player and give him additional information without changing what the other player knows. As player 1 learns something from player 2's contract offer, player 2 also learns that player 1 is better informed. As a result of all this supplementary information, not only player 1 but also player 2 may choose a different default game behavior than what would have been chosen under their prior beliefs. In the Cournot game, the change in the continuation behavior of player 2 is detrimental to player 1's payoff, even as he enjoys a higher accuracy of information. The equilibrium play of the default game under the updated information lowers player 1's payoff relative to what it would have been in the Bayesian equilibrium of the default game when every player is guided by his interim belief. The fact that the right hand side of the individual rationality constraint in (4.1) can decrease in the information revealed to player i is the key to understanding how partial information revelations enlarge the set of feasible outcome functions.[18]

[18]What is critical in this explanation is that the non-deviating player's behavior in the default game has to change depending on the information that the deviator has. In essence the equilibrium we construct punishes the deviating player by *force-feeding* him the information. If there were a way to commit the non-deviating player to a punishment strategy, there would be no need for these equilibrium path belief updates. A similar punishment could have been sustained if we did not impose a sequential rationality condition after a deviation (if we were to look for all the Bayesian equilibria rather than the *perfect* Bayesian equilibria of the reciprocal contracting game) as well.

4.2. **Incentive Compatibility.** As we argued above, the extent of the information that the players reveal with their contracts affects their continuation payoffs from a refusal to reciprocate. The potential to signal private information has an impact on how players act when they all decide to reciprocate as well. This impact can be described by the following two requirements. First, an equilibrium outcome must ensure that each player would make an accurate revelation with his contract offer in the first round. Then, once these contracts determine the mechanisms, the same outcome must give each player the incentive to reveal his true type even after observing the information leaked by the contracts in the first round.

Suppose that posterior system Π represents the equilibrium distribution of the posterior beliefs on the players' types. Each posterior in the support of Π is associated with a different *subgame* of the reciprocal contracting game, starting with the corresponding revelation in the first round. How the players' types will be mapped into their actions may vary across these subgames. Let ω^{β} be the outcome function which determines this mapping for the subgame played under posterior β.[19] We refer to the collection $\left\{\omega^{\beta}\right\}_{\beta\in supp(\Pi)}$ as a **family of outcome functions**, where $supp\,(\Pi) \subset \times_i \Delta T_i$ is the support of Π.

Once we have an outcome function for each posterior possible to reach on the equilibrium path, we can construct the outcome function associated with the overall game by taking the expectation over these posteriors. We say that the family of outcome functions $\left\{\omega^{\beta}\right\}_{\beta\in supp(\Pi)}$ is **consistent with** the outcome function ω if $\omega\,(t) = \mathbb{E}_{\beta|\Pi,t}\omega^{\beta}\,(t)$ for all type profiles t.[20] One trivial way to construct a family of outcome functions consistent with ω is setting $\omega^{\beta} = \omega$ for all $\beta \in supp\,(\Pi)$. However, as long as ω is not a deterministic outcome function and $supp\,(\Pi)$ is not singleton, one can construct other families of outcome functions consistent with ω.

In the Cournot example above, posterior system Π^* is composed of two posteriors $(\beta_1 = 0.5, \beta_2 = 0)$ and $(\beta_1 = 0.5, \beta_2 = 0.8)$, which are realized with probabilities 3/8 and 5/8 respectively. Therefore, under Π^*, a family of outcome functions consists of two functions, one for each of the two posteriors. Such a family of outcome functions is consistent with the outcome function which is constructed by taking its expectation over these two posteriors.

As mentioned above, in the first round of the reciprocal contracting game, each type of each player must have the incentive to reveal the accurate information about

[19] Notice that, given posterior β, outcome function ω^{β} maps each type profile into a randomization over actions, even when the type profile is not in the support of the posterior. That is, $\omega^{\beta}\,(t)$ is well defined even when $\beta\,(t) = 0$.

[20] For instance, if the support of the posterior system Π is finite, $\mathbb{E}_{\beta|\Pi,t}\omega^{\beta}\,(t)$ equals $\sum_{\beta\in supp(\Pi)} \Pr\left(\beta|\Pi,t\right)\omega^{\beta}\,(t)$, where $\Pr\left(\beta|\Pi,t\right) = \frac{\beta(t)\Pi(\beta)}{\beta^0(t)}$ is the conditional probability of observing posterior β given posterior system Π and type profile t.

his type. Since players decide on their revelations before they observe the revelations of the others, we refer to the conditions arising from this consideration as the pre-revelation incentive compatibility constraints. Under the posterior system Π, a family of outcome functions $\left\{\omega^{\beta}\right\}_{\beta \in supp(\Pi)}$ is **pre-revelation incentive compatible** if

$$\mathbb{E}_{\beta_{-i}|\Pi_{-i}}\mathbb{E}_{t_{-i}|\beta_{-i}}\left\{u_i\left(\omega^{\beta_i,\beta_{-i}}\left(t\right),t\right)\right\} \geq \mathbb{E}_{\beta_{-i}|\Pi_{-i}}\mathbb{E}_{t_{-i}|\beta_{-i}}\left\{u_i\left(\omega^{\beta_i',\beta_{-i}}\left(t\right),t\right)\right\} \tag{4.3}$$

for all $\beta_i, \beta_i' \in supp(\Pi_i)$ such that $\beta_i\left(t_i\right) > 0$, and for all types t_i of all players i.

Observe that pre-revelation incentive compatibility is satisfied trivially when Π does not involve any information revelation (assigns unit mass on a single distribution). Consider the Cournot example we developed above. Since Π_1^* is a degenerate distribution, the pre-revelation incentive compatibility requirement in (4.3) holds trivially for player 1. However, condition (4.3) can be rather stringent for more general distributions over posteriors. For instance, in our Cournot example, the support of Π_2^* consists of two posteriors and both these posteriors assign a non-zero probability to type (80) of player 2. Therefore any family of outcome functions satisfying the pre-revelation incentive compatibility condition must make this type indifferent between the two equilibrium path revelations he would make.

After the players offer their contracts (including the list of mechanisms δ and revelations on their types) in the first round, they have to submit their reports to the mechanisms resulting from the interaction of these contracts. In this second round of the game, the players hold additional information regarding their rivals' types, since they have already observed all the contracts. An equilibrium outcome function should give each type of each player the incentive not to imitate some other type, even under the updated equilibrium path beliefs. We capture this idea with the post-revelation incentive compatibility constraints. Under the posterior system Π, a family of outcome functions $\left\{\omega^{\beta}\right\}_{\beta \in supp(\Pi)}$ is **post-revelation incentive compatible** if for all $\beta \in supp\left(\Pi\right)$,

$$\mathbb{E}_{t_{-i}|\beta_{-i}}\left\{u_i\left(\omega^{\beta}\left(t_i,t_{-i}\right),t_i,t_{-i}\right)\right\} \geq \mathbb{E}_{t_{-i}|\beta_{-i}}\left\{u_i\left(\omega^{\beta}\left(t_i',t_{-i}\right),t_i,t_{-i}\right)\right\} \tag{4.4}$$

for each type pair t_i, t_i' of each player i.

We are now ready to suggest a definition for incentive compatibility of an outcome function. As in the case of individual rationality, this definition will refer to a specified posterior system. An outcome function ω is **incentive compatible** under the posterior system Π if there exists a family of outcome functions $\left\{\omega^{\beta}\right\}_{\beta \in supp(\Pi)}$, which is consistent with ω and which is pre-revelation and post-revelation incentive compatible under Π.

Post-revelation incentive compatibility means that each player finds it optimal to report his type truthfully whatever information the other players reveal with their contracts. An implication of this property is that committing to truthful reporting is optimal even before observing these revelations. To see this, notice that post-revelation incentive compatibility requires inequality (4.4) to hold for all revelations β in the support of the posterior system Π. After taking the expectation of both sides of this inequality over β, we end up with the following standard **interim incentive compatibility** condition

$$\mathbb{E}_{t_{-i}|\beta^0_{-i}} \{u_i(\omega(t_i, t_{-i}), t_i, t_{-i})\} \geq \mathbb{E}_{t_{-i}|\beta^0_{-i}} \{u_i(\omega(t'_i, t_{-i}), t_i, t_{-i})\} \tag{4.5}$$

for each type pair t_i, t'_i of each player i. Accordingly, if ω is incentive compatible, it is also interim incentive compatible. However, incentive compatibility is generally a more demanding condition than (4.5) since it requires truthful reporting to be optimal not only at the interim stage (under the prior β^0_{-i}), but also at the post-revelation stage (under all equilibrium path posteriors β_{-i} in the support of Π_{-i}).

4.2.1. ***Incentive Compatibility and the Cournot Default Game.*** We turn to the Cournot example one more time to demonstrate the procedure to examine incentive compatibility of an outcome function. We start with considering the outcome function which chooses the output levels that would maximize the *industry profits*, i.e., the sum of the payoffs of the two players of the Cournot game. Since the unit production costs are constant, this maximization requires that, given any type profile, the player with the higher unit cost produces zero output and the other player produces his *monopoly output*. The resulting production levels are as in the table below:

(4.6)

	$t_2 = 52$	$t_2 = 80$
$t_1 = 48$	$y_1^* = 16, y_2^* = 0$	$y_1^* = 16, y_2^* = 0$
$t_1 = 56$	$y_1^* = 0, y_2^* = 14$	$y_1^* = 12, y_2^* = 0$

In addition to deciding on their production levels, the players of the Cournot default game are allowed to make monetary transfers to each other as well. Therefore, in order to fully define an outcome function, we should also specify the type dependent transfers of the players. Now recall the reservation utilities we derived for these players in Section 4.1.1. Under refusal beliefs $\beta_i^{no} = 0$ and posterior system $\Pi^* = \{\Pi_1^*, \Pi_2^*\}$, the expected non-cooperative payoff of player 1 was 186 for type (48) and 94 for type (56). Suppose that the monetary transfers are determined in such a way that player 1 receives a payoff exactly equal to his reservation utility, regardless of the type of player 2. Below are the net transfers $\Delta z^* = z_1 - z_2$, which ensure these payoffs together with the production levels in (4.6):[21]

(4.7)

	$t_2 = 52$	$t_2 = 80$
$t_1 = 48$	Δz^*=70	Δz^*=70
$t_1 = 56$	$\Delta z^* = -94$	$\Delta z^* = 50$

We label the outcome function described in (4.6) and (4.7) as ω^*. By construction, ω^* satisfies the individual rationality constraints of player 1 under $\beta_i^{no} = 0$ and Π^*. Player 2's individual rationality constraints hold as well since his expected payoff under ω^* is $\frac{1}{2}70 + \frac{1}{2}50 = 60 \geq 0$ for type (80) and $\frac{1}{2}70 + \frac{1}{2}102 = 86 \geq 49$ for type (52).[22] Hence, outcome function ω^* is individually rational under $\beta_i^{no} = 0$ and Π^*. Observe that ω^* is interim incentive compatible as well, since it satisfies (4.5) for both players.[23]

However, ω^* does not satisfy the incentive compatibility condition we developed above. To see this, suppose that player 1 updates his belief to posterior $\beta_2 = 0$ after the first round of the reciprocal contracting game. This happens if player 2 reveals his unit cost as high (80) in the first round. In this case, any family of outcome functions which is consistent with ω^* would instruct player 2 to produce zero output regardless of the type of player 1. The reported type of player 1 determines his production level as well as the monetary transfer he will make to player 2. Consider the reporting decision of type (48) of player 1 in round 2. If this type of the player reports his type truthfully, he would produce $y_1^* = 16$, make the net transfer $\Delta z^* = 70$, and therefore receive the payoff 186, which is prescribed by the outcome function ω^*.

[21] The monopoly profit is $16^2 = 256$ under cost (48) and $12^2 = 144$ under cost (56). Therefore the net transfers reported here set the type dependent payoff of player 1 at the targeted reservation payoff. Notice that we only report the *net* transfers Δz^* since the values of z_1 and z_2 are redundant.

[22] The monopoly profit is $14^2 = 196$ under cost (52). This yields the payoff $196 - 94 = 102$ net of the transfer when player 2 type (52) faces player 1 type (56).

[23] Recall that either player has only two possible types. Interim incentive compatibility demands the difference between the expected payoffs of these two types not to be too large or too small. Otherwise one of the types would find it profitable to imitate the other one. More specifically, in our linear environment, constraint (4.5) asks for the expected payoff difference to be in-between "the difference between the unit costs of the two types" multiplied by "the expected production level of each type." For player 1, this condition can be written as

$$\begin{aligned}(56-48)\,\mathbb{E}_{t_2} y_1^*(48,t_2) &\geq 186 - 94 \geq (56-48)\,\mathbb{E}_{t_2} y_1^*(56,t_2)\\ (8)\,16 &\geq 92 \geq (8)\,\frac{1}{2}12\end{aligned}$$

given outcome function ω^*. The corresponding condition for player 2 is

$$\begin{aligned}(80-52)\,\mathbb{E}_{t_1} y_2^*(t_1,52) &\geq 86 - 60 \geq (80-52)\,\mathbb{E}_{t_1} y_2^*(t_1,80)\\ (28)\,\frac{1}{2}14 &\geq 26 \geq (28)\,0.\end{aligned}$$

Since these conditions are satisfied, ω^* is interim incentive compatible.

By contrast, if the same type imitates type (56), then he would produce a lower amount $y_1^* = 12$, make a lower net transfer $\Delta z^* = 50$, and end up with payoff $(80 - 12 - 48)\,12 - 50 = 190$, which is higher than the truthful reporting payoff 186. This example demonstrates that incentive compatibility in the reciprocal contracting environment is generally stronger than interim incentive compatibility.

In conclusion, the monopoly output levels in (4.6) are not incentive compatible together with the monetary transfers in (4.7). The resulting outcome function ω^* is *interim* incentive compatible, yet it is not incentive compatible under the posterior system Π^*, which is the only posterior system that makes this outcome individually rational. Later in Section 6, we will demonstrate the existence of a more elaborate transfer scheme which makes the monopoly output levels incentive compatible and which yields the same type dependent payoffs as in outcome ω^*. The monetary transfers of Section 6 will depend not only on the type reports submitted in round 2 of the contracting game, but also on the revelations made in round 1.

4.3. **The Characterization Theorem.** The main theorem can now be stated.

Theorem 1. *ω is a perfect Bayesian equilibrium outcome function of the reciprocal contracting game if and only if there exist refusal beliefs β^{no} and a posterior system Π under which ω is individually rational and incentive compatible.*

We prove Theorem 1 together with Proposition 1 in the Appendix. The proof consists of two parts. In the first part, we show that any perfect Bayesian equilibrium outcome function is individually rational and incentive compatible under some refusal beliefs β^{no} and some posterior system Π. This step proves the *only if* direction of the theorem. For the second part, we start with an outcome function ω which is individually rational and incentive compatible under some β^{no} and Π. We construct an equilibrium which supports the outcome function ω and which satisfies conditions (i) to (iv) of Proposition 1, proving the proposition and the *if* direction of the theorem.

Our characterization result suggests the following procedure to investigate whether an outcome function ω is supportable with a perfect Bayesian equilibrium of the reciprocal contracting game. First, find the refusal beliefs and posterior systems under which ω is individually rational. Then, examine if, for any of these posterior systems, one can construct a family of outcome functions which is consistent with ω and which satisfies the pre-revelation and post-revelation incentive compatibility conditions.

Under the degenerate posterior system which assigns unit mass on the prior β^0, our individual rationality and post-revelation incentive compatibility constraints boil down to the standard individual rationality (4.2) and interim incentive compatibility (4.5) conditions. Moreover pre-revelation incentive compatibility constraint is trivially satisfied since there is only one posterior in the support of the degenerate posterior system. This observation identifies an important subset of the equilibrium outcomes.

Corollary 1. *If outcome function* ω *satisfies conditions (4.2) and (4.5), it is a perfect Bayesian equilibrium outcome function of the reciprocal contracting game.*

Each of the outcomes identified in this corollary can be supported by an equilibrium where no information is revealed with the contract offers. This class of outcome functions also corresponds to the set of outcomes which are available through a more centralized scheme than our reciprocal contracting process. Consider a central designer who is uninformed on the types of the players and who can offer them a centralized contract. In case that this contract is accepted by all players, it regulates how they will play the default game. But if it is rejected by at least one of the players, the contract is *null and void* and the default game is played non-cooperatively. In this latter case, the designer does not have any capacity to influence the play of the default game. Conditions (4.2) and (4.5) also characterize the outcomes that this designer can implement by offering contracts which will be unanimously accepted by all types of all players.[24]

Many earlier studies of default games and contracts are based on the premise that the outcomes satisfying (4.2) and (4.5) are the only outcomes to be expected when players get together to negotiate how to play a game. As we have seen in Section 4.1 however, there are outcome functions which violate condition (4.2) and yet which are still classified as individually rational since they satisfy condition (4.1) under some non-degenerate posterior system. These outcome functions can be supported as equilibrium outcome functions of the reciprocal contracting game as long as they satisfy the incentive compatibility conditions in (4.3) and (4.4), which are generally stronger than the interim incentive compatibility constraint in (4.5).

Let us summarize our discussion on individual rationality and incentive compatibility. Due to the possibility of information revelation during the negotiation of contracts, some outcome functions qualify as individually rational even though they yield a payoff lower than what a player would expect from the play of the default game under his prior beliefs. Supporting information revelations at the contracting

[24]It is important not to confuse the *constrained* mechanism design approach here with an *unconstrained* design approach where the designer can enforce an arbitrary punishment on a non-participating player by forcing the participating players to take punitive actions they may not necessarily like and by using their private information to increase the impact of this punishment (Jehiel, Moldovanu, and Stacchetti, 1996 and 1999). Peters (2013) shows that the set of implementable outcomes by such an unconstrained mechanism designer is the same as the set of equilibrium outcomes when reciprocal contracts include punishment clauses.

stage, however, comes at the cost of more stringent incentive compatibility requirements than the standard interim conditions.

Accommodating information revelations on the equilibrium path complicates the definitions of the individual rationality and incentive compatibility requirements. Stating these conditions in our setting necessitates referring to *posterior systems* and *families of outcome functions*, which do not appear in the standard interim versions in (4.2) and (4.5). However, this does not mean that we could have obtained simpler characterization conditions if we limited the information revelation capacities of the players, say by removing the possibility of sending revelation messages as part of their reciprocal contracts. As long as a player has the ability to influence the resulting mechanism, he can use his decision as a credible signal of his private information. For example, Celik and Peters (2011) show that even a simple yes or no decision on a central designer's contract can reveal the type of the responding player and therefore extend the set of feasible outcome functions beyond what is outlined by the standard conditions. Accordingly, for conditions (4.2) and (4.5) to characterize *all* the outcomes available to negotiating parties, we need not only a central designer, but also an ad hoc directive which instructs this designer to offer only the contracts which would be acceptable by all types of all players.

5. Implications for Collusion and Mechanism Design

Elimination of non-credible punishments expands the economic applications that can benefit from the reciprocal contracting methodology. For instance, our analysis has important implications on identifying the collusion potential between multiple agents responding to a *grand contract* designed by a *principal.* If these agents cannot agree on the collusion scheme to follow, they would be released from any of the commitments they deliberated on and would respond to the grand contract non-cooperatively. The standard approach in the collusion literature is to assume that collusion is mediated by a third party, such as the above mentioned central mechanism designer.[25] This mediator aims to maximize the sum of the (ex-ante) payoffs of the colluding agents.[26]

By referring to the incentive constraints discussed in the previous section, we can formalize such a third party's optimization problem as follows. Given the default game induced by the grand contract, the third party chooses an outcome function

[25]A recent exception is the work by Zheng (2011), who refers to our reciprocal contracting approach in order to investigate collusion (formation of fund-pooling consortiums) between liquidity constrained bidders.

[26]Che and Kim (2006) consider extensions of this problem where the objective function is a weighted sum of the colluding agents' utilities with potentially different weights for different agents or for different types. They also account for the possibility that not all of the agents collude. Our discussion would apply to this general case as well.

which maximizes the sum of the ex-ante utilities of the agents subject to the standard interim individual rationality and incentive compatibility constraints in (4.2) and (4.5). The outcome which constitutes a solution to this maximization problem determines the way that the agents will respond to the grand contract.

Suppose outcome $\bar{\omega}$ is a solution to the third party's side contract selection program outlined above. A nice property of $\bar{\omega}$ is **interim incentive efficiency**: Outcome $\bar{\omega}$ is Pareto undominated within the class of interim incentive compatible outcome functions.[27] An important consequence of this observation is the **collusion proofness principle**. Instead of designing an *indirect* grand contract and making the agents collude to support outcome $\bar{\omega}$, the principal could have induced the very same outcome by designing a *direct* grand contract which asks the agents to report their types. This direct contract replicates what each type of each agent does under the outcome function $\bar{\omega}$. Suppose the agents are colluding when they play the default game induced by this direct grand contract. Since $\bar{\omega}$ is already interim incentive efficient,[28] the third party who is mediating the collusion cannot find a better course of action than suggesting the agents to reveal their types truthfully. In other words, $\bar{\omega}$ is *collusion proof.*

The collusion proofness principle, which is established by Laffont and Martimort (1997 and 2000), provides a fundamental simplification in the analysis of designing mechanisms when the responding agents have the ability to collude. Suppose we want to find the outcomes that a principal can support in this setting. Collusion proofness principle tells us that it is not necessary to consider all the grand contracts that this principal can design and the third party's reaction to each of these contracts. Instead, the result implies that the set of all the outcomes that the principal can support under collusion is identical to the set of collusion proof outcomes.

Now suppose that the collusive side contract is not designed by a hypothetical third party but it is determined through the reciprocal contracting approach developed in this paper. Unlike in the third party initiated collusion, reciprocal contracting gives the agents the chance to reveal credible information about their types during the negotiation of the side contract. As a consequence, collusion possibilities of the agents are improved under reciprocal contracting. Specifically, the characterization result in Theorem 1 asserts that the equilibrium outcomes for the reciprocal contracting

[27]To see this, suppose to the contrary that outcome function ω' is interim incentive compatible and it Pareto dominates $\bar{\omega}$. Then ω' satisfies (4.5) by hypothesis and satisfies (4.2) since it Pareto dominates $\bar{\omega}$. Moreover, ω' yields a higher total payoff than does $\bar{\omega}$, which is a contradiction to $\bar{\omega}$ being a solution to the Laffont - Martimort program.

[28]Notice that the direct grand contract is essentially constructed by removing the *irrelevant* messages – which are not sent on the equilibrium path – from the original indirect grand contract. Since $\bar{\omega}$ is interim incentive efficient under the indirect grand contract, it remains to be interim incentive efficient under the direct grand contract.

game are the individually rational and incentive compatible outcomes under arbitrary refusal beliefs and arbitrary posterior systems. As we discussed in the previous section, this set of outcomes is larger than the set of outcomes satisfying the standard interim individual rationality and incentive compatibility constraints in (4.2) and (4.5).

A surprising aspect of reciprocal contracting is that it gives the principal the opportunity to implement outcomes which are not collusion proof. In order to make the comparison between the two approaches to collusion starker, let us consider the *best* equilibrium of the reciprocal contracting game, which maximizes the sum of the ex-ante expected utilities of the agents. Suppose $\hat{\omega}$ is the outcome function resulting from this equilibrium. Given an arbitrary grand contract, both outcome $\hat{\omega}$ and outcome $\bar{\omega}$ (which is the solution to the third party's program) maximize the same objective function. However, due to the increased collusion potential under reciprocal contracting, $\hat{\omega}$ can yield a strictly higher value for the objective than does $\bar{\omega}$. In this case, $\hat{\omega}$ does not satisfy the standard individual rationality constraints in (4.2). Instead, there exists a non-degenerate posterior system $\hat{\Pi}$, under which $\hat{\omega}$ is individually rational, i.e., it satisfies (4.1). Outcome $\hat{\omega}$ also satisfies the pre-revelation and post-revelation incentive compatibility constraints in (4.3) and (4.4) under the same posterior system $\hat{\Pi}$.

This discussion reveals the possibility that outcome $\hat{\omega}$ may not be interim incentive efficient. That is, given the same grand contract, there may exist another outcome ω', which is interim incentive compatible and which Pareto dominates the "best" reciprocal contracting outcome $\hat{\omega}$. Outcome ω' fails to be a reciprocal contracting equilibrium outcome because it does not satisfy the pre-revelation and post-revelation incentive compatibility constraints in (4.3) and (4.4) under $\hat{\Pi}$, even though it is incentive compatible in the interim sense and therefore satisfies (4.5). The conclusion is that outcome $\hat{\omega}$ may not be collusion proof: suppose the principal offers a direct grand contract to implement $\hat{\omega}$. There may exist an alternative outcome function available under this direct grand contract which is interim incentive compatible and which Pareto dominates $\hat{\omega}$. In this case, the agents would agree on a reciprocal collusive contract to support this alternative outcome rather than reporting their types truthfully to support $\hat{\omega}$.[29] Consequently, outcome $\hat{\omega}$ would be available only through collusion among the agents following the design of an indirect grand contract.

This argument illustrates that collusion proofness principle does not extend to settings such as the reciprocal contracting game, where there is possibility of information revelation prior to finalizing a collusive agreement. The main problem is the

[29] Notice that this agreement does not require revelation of partial information with contract offers. Therefore the alternative outcome can be supported with a *pooling* equilibrium of the reciprocal contracting game.

discrepancy between the interim incentive compatibility constraint in (4.5) and the pre-revelation and post-revelation incentive compatibility requirements in (4.3) and (4.4). In the following section, we will show that this discrepancy vanishes once we introduce more structure for the agents' preferences.

6. Incentive Constraints under Private Values, Single Crossing, and Transferable Utility

In this section, we argue that the incentive compatibility requirements in our characterization theorem can be simplified further under the standard assumptions of private values and single crossing. In particular, we show that, for an important class of outcome functions which satisfy a monotonicity property, the incentive compatibility conditions in (4.3) and (4.4) boil down to the less demanding and more familiar interim incentive compatibility constraint in (4.5).

Player i's preferences exhibit **private values**, if his utility function depends only on the default game actions and his own type, but not on the types of the other players. In this case, the expected utility of player i can be written as $u_i(q, t_i)$, where $q \in \Delta A$ is a randomization over action profiles and $t_i \in T_i$ is player i's type. To describe the second condition we impose, we relabel types of player i such that his type space T_i is a subset of the set of real numbers $\mathbb{R}$. Preferences of player i satisfy the **single crossing property** if for any two possibly randomized action profiles q and q', $u_i(q, t_i) - u_i(q', t_i)$ is either decreasing or increasing in t_i. Single crossing property implies an *order* on the mixed action profiles. Accordingly, there exists a function (which is unique up to affine transformations) $h_i : A \to \mathbb{R}$ such that

(6.1)
$$\mathbb{E}_{a|q}\{h_i(a)\} \geq \mathbb{E}_{a|q'}\{h_i(a)\} \text{ if and only if } u_i(q, t_i) - u_i(q', t_i) \text{ is weakly decreasing in } t_i.$$

The single crossing property is trivially satisfied for players who have at most two types. Notice that, in our Cournot default game, both players' preferences exhibit private values and satisfy the single crossing property. Moreover, function $h_i(\cdot)$ which satisfies condition (6.1) can be set as the expected production level of player i. In many settings, where preferences fulfill a *one dimensional condensation condition* (Mookherjee and Reichelstein, 1992), function $h_i(\cdot)$ will have a similar natural interpretation. For instance, in independent private value auctions where each bidder's type is his own valuation, the single crossing property is satisfied when $-h_i$ equals the probability that bidder i receives the auctioned object. Similarly, in public good provision games where each provider's type is the marginal value he receives from the public good, condition (6.1) holds when $-h_i$ equals the total amount of the public good. The single crossing property allows for designing schemes which separate different types of players by assigning them different levels of h_i.

We now describe a structure which can be used by players to transfer utility among themselves in addition to coordinating their actions in the default game. We require that these utility transfers do not change the players' expected payoffs. Given a posterior system Π, we define function $x_i(t, \beta)$ as the transfer to player i when the realized type profile and posterior belief are t and β respectively. A collection of these functions $x = \{x_i\}_{i \in I}$ is a **transfer rule** if it is **budget balanced**, i.e., $\sum_i x_i(t, \beta) = 0$ for all t, β; and **outcome neutral**, i.e., $\mathbb{E}_{\beta_{-i}|\Pi_{-i}} \mathbb{E}_{t_{-i}|\beta_{-i}} \{x_i(t_i, t_{-i}, \beta_i, \beta_{-i})\} = 0$ for all t_i, β_i, and i.

Outcome neutrality implies that transfer rules do not have any effect on the incentives of players at the interim stage. Accordingly, being augmented by a transfer rule does not change the individual rationality or *interim* incentive compatibility properties of an outcome function. However, these transfers would affect a player's incentive to make the accurate revelation in round 1 of the reciprocal contacting game and his incentive to report the true type in round 2 after learning the revelations of the other players. In order to account for these effects, we revisit the definitions of pre-revelation and post-revelation incentive compatibility conditions. Under the posterior system Π, a family of outcome functions $\{\omega^\beta\}_{\beta \in supp(\Pi)}$ is **pre-revelation incentive compatible with transfer rule** x if

$$
\begin{aligned}
(6.2) \qquad & \mathbb{E}_{\beta_{-i}|\Pi_{-i}} \mathbb{E}_{t_{-i}|\beta_{-i}} \left\{u_i\left(\omega^{\beta_i, \beta_{-i}}(t), t_i\right) + x_i\left(t, \beta_i, \beta_{-i}\right)\right\} \\
\geq \; & \mathbb{E}_{\beta_{-i}|\Pi_{-i}} \mathbb{E}_{t_{-i}|\beta_{-i}} \left\{u_i\left(\omega^{\beta'_i, \beta_{-i}}(t), t_i\right) + x_i\left(t, \beta'_i, \beta_{-i}\right)\right\}
\end{aligned}
$$

for all $\beta_i, \beta'_i \in supp(\Pi_i)$ such that $\beta_i(t_i) > 0$ and for all types t_i of all players i. Similarly, under the posterior system Π, a family of outcome functions $\{\omega^\beta\}_{\beta \in supp(\Pi)}$ is **post-revelation incentive compatible with transfer rule** x if for all $\beta \in supp(\Pi)$,

$$
\begin{aligned}
(6.3) \qquad & \mathbb{E}_{t_{-i}|\beta_{-i}} \left\{u_i\left(\omega^\beta(t_i, t_{-i}), t_i\right) + x_i(t_i, t_{-i}, \beta)\right\} \\
\geq \; & \mathbb{E}_{t_{-i}|\beta_{-i}} \left\{u_i\left(\omega^\beta(t'_i, t_{-i}), t_i\right) + x_i(t'_i, t_{-i}, \beta)\right\}
\end{aligned}
$$

for each type pair t_i, t'_i of each player i.

6.1. **The Cournot Default Game with Transfer Rules.** We now consider the Cournot default game one last time to demonstrate how a transfer rule can be used in order to ensure pre-revelation and post-revelation incentive compatibility of an outcome function. With tables (4.6) and (4.7) in Section 4.2.1, we have already introduced the outcome function ω^* which satisfies the individually rationality conditions (4.1) under the refusal beliefs $\beta_i^{no} = 0$ and the posterior system Π^*. This posterior system required that player 2 reveals some partial information about his type with his equilibrium path revelations. We also showed that outcome function ω^*

satisfies the interim version of the incentive compatibility constraint in (4.5). However, we argued that ω^* is not an equilibrium outcome of the reciprocal contracting game since it fails the post-revelation incentive compatibility requirement in (4.4). Even though truth-telling is optimal in the interim stage, once player 2 reveals more information about his type, we showed that one type of player 1 would prefer to imitate the other type.

We will now augment the outcome function ω^* with a transfer rule x^* and demonstrate that the resulting outcome satisfies the pre-revelation and post-revelation incentive constraints in (6.2) and (6.3). To define x^*, we let $x_1^*(t_1, t_2, \beta_1, \beta_2)$ be a function which is constant in parameters (t_2, β_1) and which depends on parameters (t_1, β_2) as below:

$$x_1^*(48, \beta_2) = -x_1^*(56, \beta_2) = \begin{cases} 10.8 \text{ if } \beta_2 = 0 \\ -6.48 \text{ if } \beta_2 = 0.8 \end{cases}.$$

In this two player environment, the balanced budget condition is satisfied by setting $x_2^* = -x_1^*$. Recall that Π^* assigns probabilities 3/8 and 5/8 to the two posteriors $\beta_2 = 0$ and $\beta_2 = 0.8$ respectively. Hence, x^* is outcome neutral since

$$\frac{3}{8}(10.8) + \frac{5}{8}(-6.48) = 0.$$

Transfer rule x^* is constructed mainly as a means of transferring utility from one type of player 1 to the other type. Player 2 is used more like a *budget-breaker* in this construction. The direction of the transfer depends on the revelation that player 2 makes in round 1. The transfers cancel out when we take expectations over posterior β_2 or over type t_1.

Now consider the family of outcome functions $\{\omega^\beta\}_{\beta \in supp(\Pi^*)}$ such that $\omega^\beta = \omega^*$ for both posteriors in the support of Π^*. First notice that, under Π^*, outcome function ω^* is pre-revelation incentive compatible with transfer rule x^*: constraint (6.2) is trivially satisfied for player 1 since he does not make any revelation in round 1. The same condition holds as an equality for the two types of player 2 since $\mathbb{E}_{t_1}\{x_2(t_1, \beta_2)\}$ is zero regardless of his revelation. Moreover, ω^* is post-revelation incentive compatible with x^* as well: constraint (6.3) is satisfied for player 2 since ω^* is interim incentive compatible and x^* does not depend on t_2. It requires slightly more work to establish the same for player 1.[30] In Section 4.2.1, we argued that

[30] Suppose the realized posterior on the type of player 2 is $\beta_2 = 0$. Once augmented by the transfer rule x^*, outcome function w^* induces a payoff difference of $(186 + 10.8) - (94 - 10.8) = 113.6$ between the two types of player 1. Constraint (6.3) demands this difference to be bounded as below:

$$\begin{aligned} (56 - 48)\, \mathbb{E}_{t_2|\beta_2=0} y_1^*(48, t_2) &\geq 113.6 \geq (56 - 48)\, \mathbb{E}_{t_2|\beta_2=0} y_1^*(56, t_2) \\ (8)\, 16 &\geq 113.6 \geq (8)\, 12 \end{aligned}$$

outcome ω^* does not satisfy constraint (4.4) for player 1. However, now that ω^* is augmented by the transfer rule x^*, the analogous condition in (6.3) holds for both types of player 1, under both posterior distributions.

Let us summarize our findings regarding the incentive compatibility of the outcome function ω^* in the Cournot default game. ω^* is *interim* incentive compatible; yet it is not incentive compatible under the posterior system Π^*, which is the only posterior system that makes this outcome individually rational. Therefore ω^* is not an equilibrium outcome of the reciprocal contracting game. Nevertheless, ω^* satisfies the incentive compatibility conditions in (6.2) and (6.3) when it is augmented with the transfer rule x^*. This transfer rule depends on both the round 1 revelations and the round 2 type reports of the players. In the context of our Cournot default game,[31] this observation points to the existence of an equilibrium outcome function which is different from ω^* but which induces the same type dependent production levels and the same type dependent payoffs as in ω^*.

6.2. **General Default Games with Transfer Rules.** In this section, we prove a result extending what we observed in the context of the Cournot default game to general settings satisfying the private values and single crossing conditions. We show that, as long as a monotonicity property holds, an interim incentive compatible outcome function can be augmented with a transfer rule such that the resulting outcome satisfies the pre-revelation and post-revelation incentive constraints in (6.2) and (6.3). Such a transfer rule can be found for any arbitrary posterior system.

We start with recalling standard results from screening theory. Under the private values assumption and the single crossing property, incentive compatibility demands a monotonic relationship between player i's type and function $h_i(\cdot)$ which satisfies condition (6.1). In particular, interim incentive compatibility condition (4.5) implies **interim monotonicity** of the outcome function, i.e., $\mathbb{E}_{t_{-i}|\beta^0_{-i}}\{h_i[\omega(t_i,t_{-i})]\}$ is weakly decreasing in t_i for all i. For instance, in our Cournot default game example, interim monotonicity requires that each player's expected production is weakly decreasing in his cost, where the expectation is taken over the different types of the other player given the prior beliefs. Moreover, many of the incentive compatibility constraints are redundant under these conditions: if the interim (or post-revelation)

Now Suppose the realized posterior on the type of player 2 is $\beta_2 = 0.8$. Outcome w^* together with the transfer rule x^* imply a payoff difference of $(186-6.48)-(94+6.48) = 79.04$ between the two types of player 1. Constraint (6.3) demands this difference to be bounded as below:

$$\begin{aligned}(56-48)\,\mathbb{E}_{t_2|\beta_2=0.8}y_1^*(48,t_2) &\geq 79.04 \geq (56-48)\,\mathbb{E}_{t_2|\beta_2=0.8}y_1^*(56,t_2)\\ (8)\,16 &\geq 79.04 \geq (8)\left[(0.8)\,0+(0.2)\,12\right]\end{aligned}$$

Since these inequalities hold, we conclude that (6.3) is satisfied for player 1.

[31]Recall that utility transfers between the players are possible to support in this game, since it explicitly allows for monetary transfers and the payoffs are quasi-linear in these transfers.

incentive compatibility constraints are satisfied between all the "adjacent" types of player i, then all the other interim (or post-revelation) incentive compatibility constraints, including the ones between the non-adjacent types of player i, hold as well.

In order to state our result, we must also define a stronger monotonicity requirement for outcome functions. We say that outcome function ω is **ex-post monotone** if $h_i [\omega (t_i, t_{-i})]$ is weakly decreasing in t_i for all t_{-i} and all i. For instance, in our Cournot default game example, ex-post monotonicity requires that each player's production level is weakly decreasing in his cost regardless of the other player's type. Unlike interim monotonicity, ex-post monotonicity is not implied by interim incentive compatibility.

Proposition 2. *Suppose all the players' preferences exhibit private values and satisfy the single crossing condition (6.1). Suppose further that outcome function ω is interim incentive compatible and ex-post monotone. Let Π be an arbitrary posterior system. There exist a family of outcome functions $\{\omega^\beta\}_{\beta \in supp(\Pi)}$ consistent with ω and a transfer rule x such that $\{\omega^\beta\}_{\beta \in supp(\Pi)}$ is pre-revelation and post-revelation incentive compatible with x under Π.*

An interim incentive compatible outcome function can be implemented with a direct revelation mechanism where the players report their private information all at once. According to the result above, if this function is ex-post monotone as well, it can be implemented with a *gradual revelation mechanism* where the players reveal their types in two successive steps. After the first step, the players update their prior belief on the types of the other players to a posterior belief. After the second step, all the remaining private information is revealed.

Proposition 2 directly follows from Lemma 1 in Celik (2013). We provide a brief discussion of the proof here and refer the readers to that paper for the details. We guarantee consistency of $\{\omega^\beta\}_{\beta \in supp(\Pi)}$ by defining it as $\omega^\beta = \omega$ for all β. In a setting where player types are continuously distributed, the revenue equivalence theorem identifies the transfers under which the ex-post monotone outcome ω^β would satisfy the post-revelation incentive constraints in (6.3) up to a constant term. In our discrete type environment, we rely on the discrete type analogues of these transfers. We use the constant term to ensure that the resulting transfer rule is outcome neutral. By construction, $\{\omega^\beta\}_{\beta \in supp(\Pi)}$ satisfies the pre-revelation constraints in (6.2) as equalities together with the outcome neutral transfers. The next step in the proof is achieving budget balance without spoiling the incentives provided by the transfer rule. To this purpose, we follow the *expected externality* method of Arrow (1979) and d'Aspremont and Gerard-Varet (1979). However, unlike in a simultaneous revelation game, where the expectations are relevant only at the interim stage, the expectations in our two-step setting are based on the posterior distributions which are endogenously determined by the players' revelations.

Many outcome functions which are important from an economic perspective are ex-post monotone.[32] Moreover, ex-post monotonicity is sufficient but not *necessary* for the construction of $\left\{\omega^{\beta}\right\}_{\beta\in supp(\Pi)}$ and x satisfying the pre-revelation and post-revelation constraints. A weaker condition would impose monotonicity only for the types which may end up in the support of the same posterior belief. When the outcome neutrality of the transfer rule is imposed only on the equilibrium path (for t_i, β_i pairs such that $\beta_i(t_i) > 0$ instead of *all* t_i, β_i pairs), this weaker monotonicity condition is sufficient for implementation with a two-step procedure (Celik, 2013, Lemma 2).

6.3. **Collusion Proofness Revisited.** We now revisit the collusion interpretation of the reciprocal contracting procedure, which we first discussed in Section 5. We assume that the default game (which may have been induced by a grand contract) satisfies the private values and single crossing conditions. Suppose the aim is finding the "best" reciprocal contracting outcome. In Section 5, we identified this outcome as the one which maximizes the sum of the ex-ante expected utility levels of the players / agents subject to the individual rationality and incentive compatibility requirements developed in Section 4. In light of Proposition 2, we consider a *relaxed* version of this maximization problem, where the objective function and the individual rationality condition are the same as before, but the incentive compatibility requirement is replaced by the weaker *interim* incentive compatibility constraint in (4.5). We denote the solution to this relaxed problem as $\tilde{\omega}$.

Since outcome $\tilde{\omega}$ is individually rational, we know that it satisfies condition (4.1) under some refusal beliefs $\tilde{\beta}^{no}$ and posterior system $\tilde{\Pi}$. Now suppose that outcome $\tilde{\omega}$ is ex-post monotone.[33] In this case, it follows from Proposition 2 that $\tilde{\omega}$ can be augmented with a transfer rule which makes the resulting allocation satisfy the pre-revelation and post-revelation incentive compatibility conditions in (6.2) and (6.3). Assuming that the players can commit to utility transfers as part of their contracts, this observation means that the best reciprocal contracting outcome is either $\tilde{\omega}$, or a transfer rule augmented version of it inducing the same type dependent actions (without considering the utility transfers) and the same interim payoffs as $\tilde{\omega}$. Moreover, outcome $\tilde{\omega}$ is interim incentive efficient, i.e., it is Pareto undominated

[32]For instance, Mookherjee and Reichelstein (1992) show that the interim incentive compatible outcome function which maximizes the objective of a *principal* (the expected value of the gross benefit from the chosen economic alternative minus the transfers to the agents) is ex-post monotone.

[33]In many applications (such as the Cournot competition, independent private value auctions, public good provision) maximization of the (weighted) sum of players' expected payoffs yields an ex-post monotone outcome function.

within the class of interim incentive compatible outcome functions. As we argued in Section 5, this last property implies collusion proofness of $\tilde{\omega}$ as well as of its transfer rule augmented versions.

7. Conclusion

Kalai et al. (2010) propose a complete information game where each of the two players can write a contract that conditions on the other player's contract. An important feature of these conditional contracts is the punishment clauses they contain. Thanks to these clauses, a player who deviates from his equilibrium play can be punished by the other player and left with a payoff as low as his minmax value. Kalai et al. show that the equilibrium outcomes in this game coincides with the outcomes that could be supported by a centralized mechanism designer who can enforce actions of all players that agree to participate in his mechanism, including the punitive actions to be followed in case that a player unilaterally decides not to participate. Forges (2013) and Peters (2013) observe that the same characterization result stands for incomplete information games with more than two players, as long as the punishment clauses are maintained as parts of the conditional contracts.

In this paper, we consider a reciprocal contracting game of incomplete information where the punishment possibilities in the contracts are constrained by sequential rationality. We allow the players to make commitments when they all agree to some course of action. Otherwise, when there is a disagreement between the players, all contracts are *null and void*, and each player is free to choose his own action. The set of equilibrium outcomes here is much smaller than the set of outcomes described by Forges and Peters. Our reciprocal contracts support fewer outcomes than the ones available for a mechanism designer who is not constrained with the sequential rationality of the punishments.

In the absence of non-credible punishments, a better fitting benchmark for conditional contracts would be an alternative centralized design scheme where the mechanism designer is constrained not to influence the play of the players when one of them unexpectedly refuses to participate in his mechanism. This constrained design scheme is already used by the earlier literature in modeling cartel agreements between firms (Cramton and Palfrey, 1990 and 1995) and collusion between multiple agents interacting with the same principal (Laffont and Martimort, 1997 and 2000). With our characterization result, we show that the set of equilibrium outcomes of the reciprocal contracting game is larger than the outcomes supportable under this constrained mechanism design scheme. Reciprocal contracts give the players the ability to signal their private information to each other even in the event that the players cannot reach to an agreement. The constrained design approach cannot replicate this signaling aspect of reciprocal contracting.

Thinking of collusion as a reciprocal contracting game rather than a constrained design problem extends the collusion opportunities available to the agents. Almost paradoxically, the additional opportunities given to the colluding agents allows the principal to implement outcomes which would not have been implementable if collusion was mediated by a constrained mechanism designer. These supplementary outcomes are not collusion proof: their implementation requires that the principal designs an indirect grand contract and lets the agents collude on their response to it. The discrepancy between the constrained mechanism design and reciprocal contracting approaches to collusion disappears in commonly studied environments with private values, single crossing, and transferable utilities.

8. Appendix

Proof of Proposition 1 and Theorem 1.

PART I:

Suppose there exists a perfect Bayesian equilibrium of the reciprocal contracting game such that ω is the equilibrium outcome function. First, we will construct Π, β^{no}, and $\left\{\omega^{\beta}\right\}_{\beta\in supp(\Pi)}$ by using the properties of the equilibrium. Then we will show that ω satisfies the individual rationality and incentive compatibility requirements together with the constructed Π, β^{no}, and $\left\{\omega^{\beta}\right\}_{\beta\in supp(\Pi)}$.

1) Construction of Π, β^{no}, and $\left\{\omega^{\beta}\right\}_{\beta\in supp(\Pi)}$

Consider equilibrium path contracts offered by an arbitrary player i. After observing each of these contracts, other players update their beliefs on player i's type using the Bayes rule. We let the equilibrium distribution over these posteriors be Π_i. Since the ex ante expectation over the posteriors equals the prior beliefs, distribution Π_i is Bayes plausible. The posterior system Π is defined as $\{\Pi_i\}_{i\in I}$.

There are infinitely many possible mechanisms for each player and therefore there are infinitely many δ mappings from revelations to the mechanism profiles. Accordingly, whatever strategies the other players are following in equilibrium, a player can always find a list $\hat{\delta}$ which would match the lists of the other players with probability zero. Consider an arbitrary contract for player i which includes the list $\hat{\delta}$. Notice that by offering this contract, player i guarantees that the continuation game is the non-cooperative play of the default game with probability one. We let refusal belief β_i^{no} be the (possibly off the equilibrium path) posterior belief on player i's type following the observation of this contract. β^{no} equals $\{\beta_i^{no}\}_{i\in I}$.

Consider the stage of the game after the announcement of a profile of equilibrium path contracts. Consistent with the Bayes rule, the beliefs are updated to some posterior β which is in the support of Π. Starting at this stage, the perfect Bayesian equilibrium pins down the continuation strategy for each type of each player (including the types which are not in the support of the posterior, i.e., t_i such that $\beta_i(t_i) = 0$). These strategies determine a mapping from the type profiles to distributions over actions. We let ω^β be this mapping. Bayes rule implies that the family of outcome functions $\{\omega^\beta\}_{\beta \in supp(\Pi)}$ is consistent with ω.

2) Verifying the constraints

We now need to show that ω, Π, β^{no}, and $\{\omega^\beta\}_{\beta \in supp(\Pi)}$ together satisfy the individual rationality and incentive compatibility constraints.

The right hand side of the individual rationality constraint in (4.1) corresponds to the payoff from a particular (possibly off the equilibrium path) strategy for player i with type t_i. The strategy involves first offering a contract that includes the mechanism list $\hat{\delta}$ that we discussed above. This contract triggers the non-cooperative play of the default game. In the sequel, the strategy instructs player i to follow the Bayesian equilibrium strategy for the default game under the posteriors β_{-i} (which depend on the other players' contracts) and β_i^{no}. Sequential rationality requires that the other players follow their Bayesian equilibrium strategies in the continuation game as well. For the strategy explained above not to be a profitable deviation for player i with type t_i, the individual rationality constraint in (4.1) must hold.

It follows from the construction of the posterior system Π that any distribution β_i in the support of Π_i corresponds to a posterior belief on player i following the observation of an equilibrium path contract offer. Therefore, for player i with type t_i, the right hand side of the pre-revelation incentive compatibility constraint in (4.3) equals to the expected payoff from offering the contract corresponding to posterior $\beta_i' \in supp(\Pi_i)$ and then following the equilibrium continuation play.[34] For this strategy not to be a profitable deviation, condition (4.3) must be satisfied.

Similarly, post-revelation incentive compatibility condition in (4.4) follows from the fact that type t_i of player i does not strictly prefer to follow the continuation equilibrium strategy of any other type after observing the contract offers of all players.

PART II:

Suppose there exists ω which satisfies the individual rationality and incentive compatibility conditions together with some Π and β^{no}. Incentive compatibility implies existence of a family of outcome functions $\{\omega^\beta\}_{\beta \in supp(\Pi)}$ which is consistent with ω and which satisfies conditions (4.3) and (4.4). We start our proof by assuming that all outcome functions in this family are *deterministic*. That is, function ω^β maps type profiles t into a single action profile in A rather than a distribution over action

[34] In the equilibrium which supports outcome ω, there may be multiple contract offers by player i which all lead to the same posterior β_i'. In this case, the relevant deviation strategy is randomizing between these contracts with probabilities which reflect their equilibrium frequencies.

profiles, for all β in the support of Π. Under this assumption, we will construct a profile of strategies and beliefs satisfying conditions (i) - (iv) defined in Proposition 1. Then, we will argue that these strategies induce ω as the outcome function. As a third step to our proof, we will show that the strategies and beliefs we constructed constitute a perfect Bayesian Equilibrium of the reciprocal contracting game. In the final step, we will accommodate the possibility of *stochastic* ω^β by showing that the players' correlating messages can be used in order to support stochastic outcomes.

1) Strategies and Beliefs

a) Equilibrium contracts:

Recall that a contract by player i consists of a list of direct mechanisms $\delta(\cdot)$ and a revelation $\hat{\beta}_i \in \Delta T_i$. In the equilibrium we construct, all types of all players submit a unique list of mechanisms $\delta^*(\cdot)$, satisfying condition (i) of Proposition 1. We will describe this list shortly. The equilibrium also instructs each player i to make revelations only within the support of Π_i. Since Π is a belief system, Π_i is a Bayes plausible distribution of posteriors on player i's types. Therefore there exists a revelation strategy for player i where different types of this player decide on the revelations in such a way that, whenever this player makes a revelation $\hat{\beta}_i \in supp(\Pi_i)$, Bayes rule assigns the posterior $\hat{\beta}_i$ to his type. This revelation strategy is consistent with condition (ii) of Proposition 1.

On the equilibrium path, player i makes revelations only within the support of Π_i. Yet, in order to fully define function δ^*, we have to describe the values it will take for all posteriors. In our construction, whenever player i makes a revelation $\hat{\beta}_i$ which is not in the support of Π_i, the equilibrium contracts interpret this as if this player made some other revelation within the support of Π_i. To formalize this idea, we let β_i^1 be an arbitrary posterior in the support of the distribution Π_i and define a transformation function $\bar{\beta}_i : \Delta T_i \to supp(\Pi_i)$ such that

$$\bar{\beta}_i\left(\hat{\beta}_i\right) = \begin{cases} \hat{\beta}_i \text{ if } \hat{\beta}_i \in supp(\Pi_i) \\ \beta_i^1 \text{ otherwise} \end{cases} .$$

The notation $\bar{\beta}\left(\hat{\beta}\right)$ refers to the profile of posterior beliefs $\left\{\bar{\beta}_i\left(\hat{\beta}_i\right)\right\}_{i\in I}$ which is in the support of the posterior system Π.

Recall that ω^β is assumed to be deterministic for all $\beta \in supp(\Pi)$ for this step of the proof. Let $\omega_i^\beta(t) \in A_i$ be the action taken by player i in action profile $\omega^\beta(t)$. We are now ready to state the list of mechanisms that the players will submit as part of their reciprocal contracts. A direct mechanism for player i maps the type reports t and the $|I| \times K$ matrix of correlating messages n into an action in A_i. When players reveal $\hat{\beta} \in \times_{i\in I}\Delta T_i$ in the first round, function δ^* determines the mechanisms in the second round according to the following formula:

$$m_i^{\hat{\beta}}(t,n) = \omega_i^{\bar{\beta}(\hat{\beta})}(t). \tag{8.1}$$

Notice that this function is constant in n. That is, when ω^{β} is deterministic, the mechanisms do not need to depend on the correlating messages. We will later use these messages to show that a stochastic ω^{β} can also be supported by direct mechanisms as well.

b) Equilibrium beliefs:

After the first round of the game, all players observe the contract offers. Description of an equilibrium demands specifying the beliefs on each player's type as a function of the contract he offers. On the equilibrium path, player i offers contracts with the list δ^* described above and a revelation $\hat{\beta}_i$ in the support of Π_i. After observing this offer, abiding by the Bayes rule, the other players update their belief on this player's type to $\hat{\beta}_i$. Notice that these equilibrium path beliefs satisfy the "accuracy requirement" (ii) of Proposition 1. Off the equilibrium path, the beliefs on player i's type are updated to $\beta_i^1 \in supp(\Pi_i)$ when player i's contract consists of list δ^* and a revelation $\hat{\beta}_i \notin supp(\Pi_i)$; and to the refusal belief β_i^{no} when player i's contract includes a list different from δ^*.

In other words, if player i offers the list δ^* and submits a revelation $\hat{\beta}_i$ in the support of Π_i, the other players update their belief to $\hat{\beta}_i$ assuming that his revelation is "accurate." Otherwise, when he offers the list δ^* and submits a revelation outside the support of Π_i, the other players change their belief to $\beta_i^1 \in supp(\Pi_i)$. However, if the contract of player i includes a list of mechanisms other than δ^*, then the belief on this player's type is updated to the refusal belief β_i^{no} regardless of the revelation made.

c) Equilibrium reports to direct mechanisms which are induced by δ^*:

Suppose, in the first round, all players offer contracts including the list δ^* we described above. In the second round, each player i should submit a type report t_i and a correlating message vector n_i to the resulting mechanisms. In the equilibrium we construct, t_i equals the type of player i and all dimensions of n_i are uniformly and independently distributed on the set $[0,1]$, satisfying conditions (iii) and (iv) respectively in Proposition 1.

d) Off the equilibrium path default game actions:

Suppose the players' contracts do not all include the same list of mechanisms. According to the rules of the reciprocal contracting game, each player must choose a default game action in the second round. In this case, the equilibrium stipulates that each player chooses his Bayesian equilibrium action (or randomization over the actions) under the beliefs updated according to the rule (b) above.

For completeness, we should also specify the off the equilibrium path continuation strategies and beliefs for the decision nodes following the players' agreement on a list of mechanisms other than the list δ^* described above. We set these as arbitrary.[35]

2) Outcome function supported by the equilibrium

We will now argue that if players follow the strategies described above, the resulting outcome function is indeed ω. Consider the equilibrium path subgame that begins when players all submit the list δ^* and their revelations are $\hat{\beta} \in supp\,(\Pi)$. The equilibrium strategies prescribe that each player i reports his true type t_i in round 2 and determines his round 1 revelation message $\hat{\beta}_i$ in a way to support Π_i as the distribution over the posteriors on his type. The proof follows from the fact that the family of outcome functions $\left\{\omega^{\beta}\right\}_{\beta \in supp(\Pi)}$ is consistent with ω:

$$\omega\left(t\right) = \mathbb{E}_{\hat{\beta}|\Pi,t}\,\omega^{\hat{\beta}}\left(t\right) \text{ for all type profiles } t.$$

3) Sequential rationality of strategies, consistency of beliefs

In this part of the proof, we demonstrate that the strategies and beliefs described above constitute a perfect Bayesian equilibrium of the reciprocal contracting game. The behavioral strategies described in (d) are sequentially rational by construction. Beliefs in (b) are consistent with the behavior in (a), since they follow from the Bayes rule on the path of play. Strategies in (c) prescribe that players reveal their true types to the mechanisms once they update their beliefs to $\bar{\beta}\left(\hat{\beta}\right)$. Optimality of truthful revelation follows from the post-revelation incentive compatibility (4.4) of $\left\{\omega^{\beta}\right\}_{\beta \in supp(\Pi)}$. Optimality of the chosen correlating messages is trivial, since they do not affect the outcome when ω^{β} is deterministic.

We now consider deviations from the behavior described in (a). There are two types of possible deviations in the first round of the game. First, a player i with type t_i may choose to offer a contract which includes the equilibrium list δ^* together with some revelation $\hat{\beta}_i'$ such that $\hat{\beta}_i' \in supp\,(\Pi_i)$ and $\hat{\beta}_i'\left(t_i\right) = 0$.[36] Pre-revelation incentive compatibility (4.3) of $\left\{\omega^{\beta}\right\}_{\beta \in supp(\Pi)}$ implies that this is not a profitable deviation. Second, a player i with type t_i may choose to offer a contract which includes a list other than δ^*. According to the beliefs in (b), all players change their belief on player i to β_i^{no}, and the beliefs on the other players are determined by the posterior system Π. After this deviation, all players follow their non-cooperative default game actions in the second round. For player i, this continuation behavior yields an expected

[35] These decision nodes are reached only if *all* players deviate from their equilibrium behavior. Our solution concept does not impose any requirement on actions chosen on such nodes.

[36] A revelation outside of *supp* (Π_i) is strategically equivalent to the revelation β_i^1 which is in *supp* (Π_i).

payoff equal to the right hand side of the constraint (4.1). Individual rationality of the outcome function ω implies that this is not a profitable deviation either.

4) Stochastic ω^{β}

The arguments above apply when all outcome functions in family $\left\{\omega^{\beta}\right\}_{\beta \in supp(\Pi)}$ are deterministic. In this final part of the proof we discuss how the uniformly distributed correlating messages can be utilized to deal with stochastic ω^{β}, without changing the incentives provided to the players. Suppose that n_i^k is the k^{th} dimension of the correlating message sent by player i. We define n^k as $\left\lfloor \sum_i n_i^k \right\rfloor$, which is the fractional part of the real number $\sum_i n_i^k$ (or in other words $\sum_i n_i^k$ mod 1). This function aggregates the numbers $\left\{n_i^k\right\}_{i \in I}$ sent by the players into another number in the unit interval. What is crucial for our derivation is noticing that $\left\lfloor \sum_i n_i^k \right\rfloor$ is uniformly distributed on $[0,1]$ as long as all n_i^k are uniform on $[0,1]$.

When $\omega^{\bar{\beta}(\hat{\beta})}(t)$ is a stochastic function, its value indicates a probability distribution on the set of action profiles A rather than a single element of it. Recall that A is a closed subset of the Euclidean space $\mathbb{R}^K$. Therefore the value of $\omega^{\bar{\beta}(\hat{\beta})}(t)$ can be represented with a joint cumulative distribution function on $\mathbb{R}^K$. Given $\omega^{\bar{\beta}(\hat{\beta})}(t)$, let $F^k\left(\cdot|a^1,...,a^{k-1}\right)$ be the marginal cumulative distribution function for the k^{th} dimension of A conditional on $a^1,...,a^{k-1}$ being the values on the first $k-1$ dimensions.[37] Notice that $F^1,...,F^K$ are sufficient to describe the distribution $\omega^{\bar{\beta}(\hat{\beta})}(t)$. With the help of the random variables $n^1,...,n^K$, we can construct action profile $a = \left(a^1,...,a^K\right)$ iteratively as below:

$$
\begin{aligned}
a^1 &= \min\left\{v : F^1(v) \geq n^1\right\}, \\
a^2 &= \min\left\{v : F^2\left(v|a^1\right) \geq n^2\right\}, \\
&\quad .... \\
a^K &= \min\left\{v : F^K\left(v|a^1,...,a^{K-1}\right) \geq n^K\right\}.
\end{aligned}
$$

a^1 to a^K are well defined since marginal cumulative distribution functions are right continuous. Moreover the distribution of action profile a which is constructed in this way matches the distribution indicated by $\omega^{\bar{\beta}(\hat{\beta})}(t)$. This means that we can now update the direct mechanisms in (8.1), make them depend on $n^1,...,n^K$ as described as above, and generate the distribution on actions as implied by $\omega^{\bar{\beta}(\hat{\beta})}(t)$.

Notice that utilizing the correlating message in this way does not spoil incentives already provided to the players. In particular, each player is still indifferent between

---

[37] For instance, $F^1(\cdot)$ is the marginal cumulative distribution function for the first dimension of A and $F^2\left(\cdot|a^1\right)$ is the marginal cumulative distribution function for its second dimension given that a^1 is the value that the first dimension assumes.

all the correlating messages in his disposal if he expects the other players to follow a uniform distribution when choosing their own correlating messages.[38] □

References

[1] Arrow, K., 1979. "The property rights doctrine and demand revelation under incomplete information" in Economics and Human Welfare, edited by M. Boskin. New York: Academic Press.

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

[3] Bachi, B., S. Ghosh, and Z. Neeman, 2013. "Communication and deception in 2-player games" Working paper.

[4] Caillaud, B. and P. Jehiel, 1998. "Collusion in auctions with externalities" Rand Journal of Economics, 29, 680-702.

[5] Celik G., 2013. "Sense and suspense: implementation by gradual revelation" Working paper.

[6] Celik G. and M. Peters, 2011. "Equilibrium rejection of a mechanism" Games and Economic Behavior, 73, 375-387.

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

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

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

[10] d'Aspremont, C. and L. A. Gerard-Varet, 1979. "Incentives and incomplete information" Journal of Public Economics, 11, 25-45.

[11] Forges, F., 2013. "A folk theorem for Bayesian games with commitment" Games and Economic Behavior, 78, 64-71.

[12] Fudenberg, D. and J. Tirole, 1991. "Perfect Bayesian equilibrium and sequential equilibrium" Journal of Economic Theory, 53, 236-260.

[13] Hagenbach, J., F. Koessler, and E. Perez-Richet, 2014. "Certifiable pre-play communication: Full disclosure" Econometrica, 82, 1093-1131.

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

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

[16] Kalai, A. T., E. Kalai, E. Lehrer, and D. Samet, 2010. "A commitment folk theorem" Games and Economic Behavior, 69, 127-137.

[38] Suppose player i expects that all the other players choose the k^{th} dimension of their correlating messages with respect to the uniform distribution independently of their types, beliefs, revelations, and reports. Therefore he believes that $\left\lfloor \sum_{j\neq i} n_j^k \right\rfloor$ is uniformly distributed on $[0,1]$. Accordingly, for all $n_i^k \in [0,1]$, random variable $\left\lfloor n_i^k + \sum_{j\neq i} n_j^k \right\rfloor$ has a uniform distribution on $[0,1]$ as well. This proves that player i is indifferent between all the correlating messages in his disposal. In other words, provided that a player believes the others are doing the same thing, it is a best reply for him to select his correlating message using a uniform distribution.

[17] Kamenica, E. and M. Gentzkow, 2011. "Bayesian persuasion" American Economic Review, 101, 2590-2615.
[18] Koessler, F. and A. Lambert-Mogiliansky, 2013. "Committing to transparency to resist corruption" Journal of Development Economics, 100, 117-126.
[19] Laffont, J. J. and D. Martimort, 1997. "Collusion under asymmetric information" Econometrica, 65, 875-911.
[20] Laffont, J. J. and D. Martimort, 2000. "Mechanism design with collusion and correlation" Econometrica, 68, 309-342.
[21] Maskin, E. and J. Tirole, 1990. "The principal - agent relationship with an informed principal: The case of private values" Econometrica, 58, 379-409.
[22] Maskin, E. and J. Tirole, 1992. "The principal - agent relationship with an informed principal II: Common values" Econometrica, 60, 1-42.
[23] Mookherjee, D. and S. Reichelstein, 1992. "Dominant strategy implementation of Bayesian incentive compatible allocation rules" Journal of Economic Theory, 56, 378 - 399.
[24] Myerson, R. B., 1991. "Game theory: Analysis of conflict" Harvard University Press: Cambridge MA, London England.
[25] Peters, M., 2013. "Reciprocal Contracting" Discussion paper, University of British Columbia.
[26] Peters, M. and B. Szentes, 2012. "Definable and Contractible Contracts," Econometrica, 80, 363-411.
[27] Peters, M. and C. Troncoso-Valverde, 2013. "A Folk Theorem for Competing Mechanisms," Journal of Economic Theory, 148, 953-973.
[28] Philippon, T. and V. Skreta, 2012. "Optimal interventions in markets with adverse selection" American Economic Review, 102, 1-29.
[29] 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.
[30] Tennenholtz, M., 2004. "Program Equilibrium," Games and Economic Behavior, 49, 363–373.
[31] Tirole, J., 2012. "Overcoming adverse selection: How public intervention can restore market functioning" American Economic Review, 102, 29-59.
[32] Zheng, C., 2011. "Polarization in competitive bidding" Working Paper.