# RECIPROCAL RELATIONSHIPS AND MECHANISM DESIGN

GORKEM CELIK AND MICHAEL PETERS

ABSTRACT. We study an incomplete information game in which players are involved in a reciprocal relationship that allows them to coordinate their actions by contracting among themselves. We model this as a competing mechanism game in which players have the ability to write contracts. We characterize the set of outcome functions that can be supported as equilibrium in this enhanced game. We use our characterization to show that the set of supportable outcomes is bigger than the set of outcomes supported by a centralized mechanism designer who can offer mechanisms in which all players participate. The difference is that the contracting game makes it possible for players to convey partial information about their type at the time they offer contracts.

It is common to use ideas from mechanism design, a tool of normative economic theory, to try to describe what is attainable by players in a game of incomplete information in which some players can contract. For example, many of the best known models of collusion assume that colluding players have access to a disinterested coordinator who collects information and enforces outcomes on their behalf.[1] Implicit in this approach is the idea that colluding players are acting cooperatively. To assume a disinterested designer can coordinate all actions is a natural extension of the kind of analysis done for games of complete information (for example Farell and Shapiro, 1990) in which cooperating players (a cartel in their case) were simply assumed to maximize joint profits.

However, a more natural approach is to model collusion itself as a non-cooperative game of incomplete information in which 'colluding' parties are simply non-cooperative

*Date*: August 1, 2011.
PRELIMINARY AND INCOMPLETE

Celik: ESSEC Business School and THEMA Research Center, celik@essec.fr ;
Peters: University of British Columbia, peters@econ.ubc.ca.

[1]Examples include Laffont and Martimort (1997), Che and Kim (2006).

players who have some partial ability to communicate with other players and to make commitments based on these communications.[2] The complication in doing this is that the extensive form game that is played in order to arrive at the cooperative outcome is usually unknown. When some players are colluding, often an illegal activity, players will go to great lengths to hide the way they negotiate and enforce an agreement.

A descriptive model of the negotiation process is likely to be complicated since it needs to capture complex strategic interactions among incompletely informed players. Furthermore, it will have to make arbitrary decisions about bargaining power and timing. Equilibrium outcome functions are unlikely to be robust with respect to these choices.

To get around this, we use the reciprocal contracting method to characterize the set of outcome functions that can be supported as equilibria in competing contract games. Reciprocal contracts condition outcomes directly on contracts of the other players. However, they do this in a very simple way. If the other contracts 'agree' with one another, then the contract implements some kind of cooperative action. If they don't 'agree', a reciprocal contract implements a punishment. We describe this method in detail below. It is shown in Peters (2010) that an outcome function can be supported as a Bayesian equilibrium in *some* extensive form game of competing contracts if and only if it can be supported as an equilibrium in a reciprocal contracting game. In this sense, reciprocal contracts provide a kind of reduced form method for studying what amounts to a wide variety of different negotiation games. This is ideal for collusive problems where it is so hard to know exactly how players are negotiating.

However, the equivalence of contracting games and reciprocal contracting games applies only to Bayesian equilibrium. We are interested in a much smaller set of equilibrium outcomes here. In particular, we want outcomes to be sequentially rational, so we want equilibrium that satisfies a refinement. Furthermore, the nature of collusion is such that it cannot really be sustained unless all players agree to it. If participating players are hurt by a punishment the others try to impose, they may well have recourse to legal action to prevent the punishment. For this reason, we

[2]We model this as a static game and make no distinction between contracts that are legally enforceable, and those that are enforced by repeated interaction.

believe that outcomes should be such that players simply play some noncooperative game unless they can come to some unanimous agreement.

Our main objective is to show that players who contract non-cooperatively can actually do more than the disinterested coordinator who typically guides collusion in the literature. The reason is that non-cooperative games give players a chance to convey information that can be used to change incentive compatibility and individual rationality constraints. The reciprocal contracting method provides a relatively straightforward way to do this. We characterize the set of perfect Bayesian equilibrium outcomes for a reciprocal contracting game with restricted punishments and use this to show how these outcomes can dominate the set supportable by a disinterested coordinator.

## 1. The Model

Our basic approach is to model a game of competing mechanisms in which players have some limited ability to condition their actions on different messages that are sent during the game. Rather than trying to model the extensive form contracting game explicitly, we adopt the reciprocal contracting approach in Peters (2010) and assume that players offer contracts that specify some kind of cooperative actions provided other players' contracts do the same thing. In the event that players contracts don't 'do the same thing', we assume that contracts are limited to a punishment that involves playing some default game non-cooperatively. The logic of this approach is based on a theorem in Peters (2010) which shows that every outcome function supported as a Bayesian equilibrium in some sequential game in competing mechanisms can be supported as an equilibrium in reciprocal contracts. The advantage of this approach is that contracts are based on a simple 'cooperate or be punished' logic that makes it very easy to characterize supportable outcomes.

We impose two restrictions on these contracts by assuming that punishments can be no more complex than playing the default game non-cooperatively. In addition, we restrict attention to outcome functions that can be supported as perfect Bayesian equilibrium. The default game might be an auction, or perhaps a Cournot competition (possibly even a repeated Cournot game).

A simple analogy is to the case where a pair of firms either comes to a mutually binding agreement, or doesn't. If there is no agreement the default game is played. Our approach goes far beyond this, however, since we need outcomes to depend on types and to convey information that will be important in the case that there is no agreement.

Secondly, we are going to require that the way players expect everyone to play the default game after a deviation in the reciprocal contracting game constitutes a Bayesian equilibrium for the corresponding continuation game whenever this is possible.[3] In order to retain the idea that sensible off equilibrium beliefs about a player can only change after that player himself has deviated, we will assume from the outset that types are independently distributed.

1.1. **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 an action set $A_i$ and a finite type set $T_i$. In standard notation $A$ and $T$ are cross product spaces representing all players' actions and types respectively. Player $i$'s type is distributed with respect to the prior distribution $\beta_i^0$, independently of the types of the other players. We define $\beta^0 = \left\{\beta_i^0\right\}_{i \in I}$ as the collection of these priors. $\beta_i^0\left(t_i\right)$ is the probability that player $i$ has type $t_i \in T_i$ under the prior distribution. Similarly, $\beta^0\left(t\right) = \prod_{i \in I} \beta_i^0\left(t_i\right)$ is the probability that the realization of the type profile is $t = \left\{t_i\right\}_{i \in I} \in T$.

Preferences of player $i$ are given by the payoff function $u_i : A \times T \rightarrow \mathbb{R}$. Players have expected utility preferences over lotteries. If $q$ is a randomization over action profiles ($q \in \Delta A$) and $t$ is a type profile, then $u_i\left(q, t\right)$ 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 $\omega : T \rightarrow \Delta A$.

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 $\beta^0 = \left\{\beta_i^0\right\}_{i \in I}$. However, 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 $\beta = \left\{\beta_i\right\}_{i \in I}$, rather

[3] We explain in more detail below why we need this 'whenever this is possible' qualification.

than the prior distribution. As in the definition of the prior distribution, $\beta_i$ is an element of $\Delta T_i$ 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 $\beta$ as the default game under belief $\beta$. 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 $\beta$ if any action in the support of randomization $q_i(t_i|\beta)$ is a solution to

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

for all types $t_i \in T_i$ of all players $i \in I$. The operator $\mathbb{E}_{t_{-i}|\beta_{-i}}$ stands for the expectation over the values of $t_{-i}$ given belief $\beta_{-i}$.[4]

We restrict attention to default games for which a Bayesian equilibrium exists. Recall that the type sets are finite and therefore the 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.

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

$$U_i(t_i, \beta) = \mathbb{E}_{t_{-i}|\beta_{-i}}\left[u_i\left(q_i\left(t_i|\beta\right), q_{-i}\left(t_{-i}|\beta\right), t_i, t_{-i}\right)\right]. \tag{1.1}$$

1.2. **Example: The Cournot 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 0.6 or 0.7. Unit cost of player 2 is either 0.65 or 1. The inverse demand function for the good they produce is given

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

as $P = 1 - (y_1 + y_2)$, 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 $u_i(y_i, y_j, t_i) = [1 - (y_i + y_j) - t_i]\, y_i$. Suppose that types of the players are independently distributed. 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 $\beta_1$ and $\beta_2$ be the probabilities that player 1 has type 0.6 and player 2 has type 0.65 respectively. There is a unique Bayesian equilibrium of this game under any pair of beliefs $(\beta_1, \beta_2)$. The resulting equilibrium production and expected payoff levels are as in the table below:

| | |
|---|---|
| $y_1(0.7\|\beta_1, \beta_2) = \frac{0.6-0.35\beta_2+0.05\beta_1\beta_2}{4-\beta_2}$ | $U_1(0.7, \beta_1, \beta_2) = \left(\frac{0.6-0.35\beta_2+0.05\beta_1\beta_2}{4-\beta_2}\right)^2$ |
| $y_1(0.6\|\beta_1, \beta_2) = \frac{0.8-0.4\beta_2+0.05\beta_1\beta_2}{4-\beta_2}$ | $U_1(0.6, \beta_1, \beta_2) = \left(\frac{0.8-0.4\beta_2+0.05\beta_1\beta_2}{4-\beta_2}\right)^2$ |
| $y_2(1\|\beta_1, \beta_2) = 0$ | $U_2(1, \beta_1, \beta_2) = 0$ |
| $y_2(0.65\|\beta_1, \beta_2) = \frac{0.4-0.1\beta_1}{4-\beta_2}$ | $U_2(0.65, \beta_1, \beta_2) = \left(\frac{0.4-0.1\beta_1}{4-\beta_2}\right)^2$ |

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 0.6 and 0.7 of player 1 and type 0.65 of player 2), the equilibrium behavior and the expected payoff depend on the beliefs under which the game is played. As expected, player $i$'s Bayesian equilibrium output and expected payoff are weakly decreasing in the belief that his rival (player $j$) has the lower cost ($\beta_j$), and weakly increasing in player $j$'s belief that player $i$ has the lower cost ($\beta_i$).

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, each player can also make a non-negative transfer to the other player. Each player maximizes his utility net of the transfers.[5] 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 version of the game as well. This game, which we call the

[5] Player $i$'s utility net of the transfers is $\tilde{u}_i(y_i, y_j, z_i, z_j, t_i) = [1 - (y_i + y_j) - t_i]\, y_i - z_i + z_j$, where $z_i$ and $z_j$ are the monetary transfers made by firms $i$ and $j$ respectively.

Cournot game, is a slightly modified version of the example studied by Celik and Peters (forthcoming). In what follows, we will refer to the Cournot game to illustrate the key points of our analysis.

## 2. The Contracting Game

The contracting process takes place in two rounds. In the first round, players offer *contracts*. These offers determine a *mechanism* for each player, which commits this player to an action contingent on messages that are 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 (2008), or implicit as in the contracting game we explain below.

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 direct mechanism for player $i$ is a mapping from a collection of $2|I|$ arguments into the player's action:

$$m_i : T \times [0,1]^{|I|} \to A_i.$$

The first $|I|$ arguments represent the type reports made by players in the second round. The following $|I|$ arguments (real numbers from the unit interval $[0,1]$) would be less familiar to most readers. These numbers are the *correlating messages* which are submitted by the players as well. 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 messages can be used in order to support randomizations over actions. Moreover we will show that these randomizations may be correlated across the players as well.[6] The set of direct mechanisms for player $i$ is defined to be $M_i$.

[6]In other words, we use the correlating messages 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 $i$.

In the first round of our 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 $\beta_i^0$ governs the distribution of player $i$'s types. After observing the revelation made by player $i$, the prior belief on this player may be updated to a posterior belief. We need the set of possible revelations by each player to be rich enough to support any possible posterior distribution. For this reason, we model each player's revelation as announcing a distribution of his types.[7] Formally, a reciprocal contract for player $i$ consists of a revelation $\hat{\beta}_i \in \Delta T_i$ and a list of potential direct mechanisms $\delta$ which is represented by a mapping from revelations of all players into profiles of direct mechanisms

$$\delta : \times_i \Delta T_i \to \times_i M_i.$$

These contracts determine the players' mechanisms as follows: If all contracts include the same list $\delta$, then the mechanisms are indeed pinned down by how this function maps the submitted revelations of players into a mechanism profile. That is, if the players' offers in the first round agree on function $\delta$, than the direct mechanism $\delta_i(\beta)$ determines the mechanism that player $i$ will follow in the second round, where $\delta_i$ is the $i^{th}$ component of function $\delta$. However, if there is at least one player who offered a contract containing a different list $\delta$ than did the other players, then 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.[8]

[7]This reasoning assumes that all players other than player $i$ would update their beliefs in the same way after observing player $i$'s revelation. This assumption is consistent with Fudenberg and Tirole's (1990) definition of perfect Bayesian equilibrium.

[8]In the setting we study, when an "agreement" fails, the harshest punishment that can be imposed on the player who caused things to break down is the non-cooperative play of the default game. By contrast, Peters and Szentes (2008) allow for more prohibitive punishments (in the form of restrictions on the available actions for the non-deviating players). They prove an "invariant

A contract offer, by construction, leads to a specific commitment for a player. Yet his contract 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 $\delta$ 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 consist of strategies and beliefs satisfying the conditions below:

i) In round 1, each type of each player $i$ chooses the contract(s) 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.

iii) In round 2, on the equilibrium path or on the continuation games reachable by unilateral deviations from the equilibrium play[9], each type of each player $i$ chooses his message to the mechanisms (if all contract offers include the same list of direct mechanisms) or his default game action to maximize his expected continuation payoff, given the updated beliefs.

## 3. Incentive Constraints

The main objective of this paper is providing a characterization of the outcome functions which are supportable as the Perfect Bayesian equilibrium outcomes of the

punishment property," which roughly says that all different deviations by the same player must be punished the same way.

[9]Our definition of perfect Bayesian equilibrium demands for sequential rationality of strategies only for continuation games which are either on the equilibrium path or accessible by unilateral deviations from the equilibrium. As Peters and Troncoso Valverde (2010) 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.

reciprocal contracting game. With our first result, we show that it is sufficient to restrict attention to a specific class of equilibria for this task.

**Proposition 1.** *If an outcome function $\omega$ is supportable as a perfect Bayesian equilibrium of the reciprocal contracting game then it is also supportable by a perfect Bayesian equilibrium of this game where*

*i) players reciprocate: all types of all players submit a unique list of mechanisms $\delta^*$ as part of their contracts in round 1;*

*ii) revelations are accurate: on the equilibrium path, after observing the 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 messages are uniformly distributed: on the equilibrium path, each player's correlating message is uniformly distributed on the interval $[0,1]$ regardless of his type and the posterior beliefs.*

We provide the proof of this proposition after we state Theorem 1 at the end of this section.

Property (i) above follows from a familiar argument. Suppose there exists an equilibrium where some types of some players do not reciprocate: They submit a list of mechanisms other than $\delta^*$. The resulting equilibrium outcome can be 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 as public randomization devices in this class of equilibria.

The intuition for property (ii) follows from the revelation principle as well. If all types of all players submit the same list of mechanisms $\delta^*$, 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 by refusing to reciprocate with the others players, 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 conditions below and discuss how they relate to the more familiar versions of incentive constraints. Then we provide a formal characterization of equilibrium outcome functions referring to the conditions we developed.

3.1. **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, players receive their non-cooperative payoffs defined in (1.1). For this deviation not to be profitable, each type of each player must expect a weakly higher payoff from the equilibrium outcome than their non-cooperative payoff. This consideration will yield the individual rationality constraints.

After a player's refusal to reciprocate, the beliefs on the types of the players need not remain the same as the prior beliefs. First, as a result of the refusal of player $i$, the other players may update their belief regarding the type of this player from prior $\beta_i^0$ to some posterior $\beta_i^{no}$.[10] 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.[11]

[10]It is implicit in this notation that when player $i$ takes an action (on or off the path of play), players $j$ and $k$ update their beliefs on player $i$ in the same way.

[11]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 rejection beliefs consist 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 and Tan and

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 reciprocating players as well. Recall that players are allowed to make revelations about their private information as part of their contract offers. As a result, their equilibrium play may indeed unveil information on their types. When continuation play has to satisfy a refinement like *perfect* Bayesian equilibrium, 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 $\beta_i^{no}$ if he does not reciprocate. He also comprehends that after observing the contract offers, the beliefs on the other players' types will also be updated to some posterior $\beta_{-i}$. Recall that the non-cooperative payoff $U_i\left(t_i, \beta_i^{no}, \beta_{-i}\right)$ defined in (1.1) 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 observes the other contracts, at a time that he does not know the exact realization of the posterior $\beta_{-i}$. 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 $\Pi_i \in \Delta\left(\Delta T_i\right)$. 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 to the prior distribution: $\mathbb{E}_{\beta_i|\Pi_i}\beta_i = \beta_i^0$. Following Kamenica and Gentzkow (forthcoming), we call distribution $\Pi_i$ **Bayes plausible** when it satisfies this property. If $\Pi_i$ is Bayes plausible for each player $i$, then we refer to the collection $\Pi = \left\{\Pi_i\right\}_{i\in I}$ as a **posterior system.**

Suppose the outcome function $\omega$ 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 unilaterally deviate by refusing to reciprocate. This incentive is represented by the following condition. Under the refusal beliefs

Yilankaya, 2007). We will continue our analysis without imposing such a refinement and allowing for arbitrary rejection belief.

$\beta^{no}$ and the posteriors system $\Pi$, outcome function $\omega$ is **individually rational** if

$$\mathbb{E}_{t_{-i}|\beta^0_{-i}}\left\{u_i\left(\omega\left(t\right),t\right)\right\} \geq \mathbb{E}_{\beta_{-i}|\Pi_{-i}}\left\{U_i\left(t_i,\beta_i^{no},\beta_{-i}\right)\right\} \tag{3.1}$$

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 $\Pi$ which puts unit mass on the prior distribution $\beta^0$.

Under this system, the right hand side of (3.1) boils down to $U_i\left(t_i,\beta_i^{no},\beta^0_{-i}\right)$:

$$\mathbb{E}_{t_{-i}|\beta^0_{-i}}\left\{u_i\left(\omega\left(t\right),t\right)\right\} \geq U_i\left(t_i,\beta_i^{no},\beta^0_{-i}\right) \tag{3.2}$$

for all $t_i$ and all $i$.

For player $i$, $U_i\left(t_i,\beta_i^{no},\beta^0_{-i}\right)$ is the non-cooperative default game payoff corresponding to the lowest level of information. Any other posterior system would have given player $i$ more information on the types of the other players. Nevertheless, the following discussion demonstrates that $U_i\left(t_i,\beta_i^{no},\beta^0_{-i}\right)$ does not necessarily constitute a lower bound on the right hand side of (3.1). In other words, the equilibrium payoff of player $i$ may decrease with the level of information revealed by the contracts offered in the first round.[12]

3.1.1. ***Individual Rationality and the Cournot Game.*** Suppose that the two players of the Cournot game introduced in the previous section are negotiating over a cartel agreement by using the reciprocal contracting process we described. If they can agree on it, the cartel agreement will determine the production levels of the players and the monetary transfers they will make to each other. Each player can refuse to reciprocate with the other player and trigger the non-cooperative play of the Cournot game. What payoff would a player expect from such a deviation? Player 2 with cost 1 receives zero payoff from the non-cooperative play of the default game regardless of the beliefs. For the other types of players, the payoff functions described in the previous section imply that player $i$ would make the lowest non-cooperative payoff

[12] It is possible as long as the default game payoff $U_i$ is not convex in the posterior belief over the rivals' types $\beta_{-i}$. Celik and Peters (forthcoming) show that even a simple two player Cournot game with linear demand and constant unit cost can support a non-convex payoff function.

when $\beta_i = 0$. This means that a larger set of outcome functions will be classified as individually rational under the refusal beliefs $\beta_i^{no} = 0$.

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

Let us start with considering the non-cooperative payoff function of player 2 with cost 0.65. This type's non-cooperative payoff $U_2(0.65, \beta_1, \beta_2)$ is convex in $\beta_1$, implying that his expected payoff would increase in the information he receives on player 1's type. Therefore the right hand side of (3.1) would be minimized if distribution $\Pi_1^*$ assigns unit mass on the prior distribution $\beta_1^0 = 0.5$. Under the refusal belief $\beta_2^{no} = 0$ and the prior belief $\beta_1^0 = 0.5$, the non-cooperative payoff for type 0.65 of player 2 is $U_2(0.65, 0.5, 0) = \left(\frac{0.4-0.05}{4}\right)^2 = 49/6400$. Player 2 with type 0.65 will not accept the cartel agreement unless he receives a payoff lower than this figure.

Now we turn our attention to player 1. Regardless of his type, this player's non-cooperative payoff is not convex in $\beta_2$. This non-convexity indicates that expected payoff of player 1 can be reduced by revealing information to him on player 2's type. Concentrate on player 1 with type 0.7 for now. Suppose first 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 for type 0.7 of player 1 is $U_1(0.7, 0, 0.5) = \left(\frac{0.6-0.175}{3.5}\right)^2 = 289/19600$. If player 1 with type 0.7 believes that he will not receive any additional information about his rival, this is the lowest payoff he will agree on during the contracting process. However, thanks to the non-convexity of function $U_1$, one could reduce the reservation payoff of the same player by revealing information about the cost of the rival player. There exist non-degenerate distributions of posteriors $\Pi_2$ which would reduce the expected value of $U_1$ while respecting the Bayes plausibility condition $\mathbb{E}_{\beta_2|\Pi_2}\beta_2 = 0.5$. Celik and Peters (forthcoming) show that the expected non-cooperative payoff of player 1 with type 0.7 is minimized under the distribution $\Pi_2^*$ which assigns probability 3/8 to posterior $\beta_2 = 0$ and probability 5/8 to posterior $\beta_2 = 0.8$. In this case, $\mathbb{E}_{\beta_2|\Pi_2^*}U_1(0.7, 0, \beta_2) = 47/3200$ which is lower than 289/19600. We can also compute the expected non-cooperative payoff of type 0.6 under distribution $\Pi_2^*$ as $\mathbb{E}_{\beta_2|\Pi_2^*} U_1(0.6, 0, \beta_2) = 93/3200$.

How can a player's payoff decrease in the level of information that is revealed to this player? The answer lies in the observation that it is not possible to single out one player and give him additional information without changing what the other players know. 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 additional information, not only player 1 but also player 2 may choose different default game actions than they would have done 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 be in the Bayesian equilibrium of the default game when every player is guided by his or her interim beliefs. The fact that the right hand side of the individual rationality constraints (3.1) can decrease in the information revealed to player $i$ is the key to how partial information revelation at the contracting stage enlarges the set of feasible outcome functions.[13]

This example also illustrates how reciprocal contracting can achieve more than what an uninformed designer can accomplish by offering a centralized contract unanimously acceptable to all types of all players. There is no possibility of information revelation when ratifying a designer-offered contract. Consequently, condition (3.2) would be necessary for unanimous acceptability of a such a contract. In the Cournot game, player 1 with type 0.7 would not accept a designer-offered contract which yields a payoff lower than 289/19600. However, reciprocal contracts support lower payoffs for this type of player by allowing for equilibrium path signaling.

[13]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, 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.

We now defined two distributions over the posteriors on the types of player 1 ($\Pi_1^*$) and player 2 ($\Pi_2^*$). These two constitute a posterior system $\Pi^* = \{\Pi_1^*, \Pi_2^*\}$. And our discussion above reveals the values on the right hand sides of constraints in (3.1) under this system.

3.2. **Incentive Compatibility.** As we argued above, the extent of the information that the players reveal with their contracts affects the continuation payoff from refusing 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 function must ensure that each player would make the accurate revelation with his contract offer in the first round. Then, once these contracts determine the mechanisms, the same outcome function must give each player the incentive to reveal his true type even after observing the information leaked by the contracts in the first round.

We start our discussion by defining an *extended outcome function*. This function accommodates the possibility that actions taken in the default game may depend on the information revealed with contract offers. Given a posterior system $\Pi$, an extended outcome function is defined as

$$\omega_\Pi : T \times supp\,(\Pi) \to \Delta A,$$

where $supp\,(\Pi) \subset \times_i \Delta T_i$ is the support of $\Pi$.[14] Taking the expectation of the extended outcome function $\omega_\Pi$ over different posteriors yields an outcome function $\omega$. We say that $\omega_\Pi$ is **consistent with** $\omega$ if $\omega\,(t) = \mathbb{E}_{\beta|\Pi}\omega_\Pi\,(t, \beta)$. Function $\omega_\Pi$ is the **natural extension** of the outcome function $\omega$ if it prescribes the same randomization over the action profiles as $\omega$, regardless of the realized posterior, i.e, if $\omega_\Pi\,(t, \beta) = \omega\,(t)$ for all $\beta \in supp\,(\Pi)$. The natural extension of an outcome function is always consistent with it. However one can construct other consistent extended outcome functions.

In the Cournot example above, posterior system $\Pi^*$ consists 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

[14]Notice that, given posterior $\beta$, the extended outcome function $\omega_\Pi$ maps each type profile into a randomization over actions, even when the type profile is not in the support of the posterior. That is, $\omega_\Pi\,(t, \beta)$ is well defined even when $\beta\,(t) = 0$.

and 5/8 respectively. An extended outcome function can assign different type dependent actions (or randomizations over actions) to these two posteriors. Such an extended outcome function 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 his type. Since each player is deciding his revelation before he sees 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 $\Pi$, an extended outcome function $\omega_\Pi$ is **pre revelation incentive compatible** if

(3.3)

$$\mathbb{E}_{\beta_{-i}|\Pi_{-i}}\mathbb{E}_{t_{-i}|\beta_{-i}}\left\{u_i\left(\omega_\Pi\left(t,\beta_i,\beta_{-i}\right),t\right)\right\} \geq \mathbb{E}_{\beta_{-i}|\Pi_{-i}}\mathbb{E}_{t_{-i}|\beta_{-i}}\left\{u_i\left(\omega_\Pi\left(t,\beta_i',\beta_{-i}\right),t\right)\right\}$$

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 $i$.

Observe that the pre revelation incentive compatibility is trivially satisfied when $\Pi$ does not involve any information revelation (assigns unit mass on a single distribution). Consider the Cournot example we developed above. Since $\Pi_1^*$ is a degenerate distribution, pre revelation incentive compatibility (3.3) is trivially satisfied for player 1. By contrast, for *generic* distributions over posteriors, (3.3) is a rather stringent condition. In our Cournot example, the support of $\Pi_2^*$ consists of two posteriors and both these posteriors assign a non-zero probability to type 1 of player 2. Therefore an extended outcome function satisfying pre-revelation incentive compatibility condition should make this type indifferent between the expected continuation payoffs following either one of these posteriors.

After the players offer their contracts (including the list of mechanisms $\delta$ 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 the types of their rivals, since they have already observed their 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 constraints. Under the posterior system $\Pi$, an extended outcome function $\omega_\Pi$ is **post revelation incentive compatible** if for all $\beta \in supp\,(\Pi)$,

$$\mathbb{E}_{t_{-i}|\beta_{-i}}\left\{u_i\left(\omega_\Pi\left(t_i, t_{-i}, \beta\right), t_i, t_{-i}\right)\right\} \geq \mathbb{E}_{t_{-i}|\beta_{-i}}\left\{u_i\left(\omega_\Pi\left(t_i', t_{-i}, \beta\right), t_i, t_{-i}\right)\right\} \tag{3.4}$$

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 $\omega$ is **incentive compatible** under the posterior system $\Pi$ if there exists an extended outcome function $\omega_\Pi$, which is consistent with $\omega$ and which is pre revelation and post revelation incentive compatible under $\Pi$.

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 truthful reporting is optimal even before observing these revelations. To see this, notice that post revelation incentive compatibility requires inequality (3.4) to hold for all revelations $\beta$ in the support of the posterior system $\Pi$. After taking the expectation of both sides of this inequality over $\beta$, we end up with the following standard *interim incentive compatibility* condition:

$$\mathbb{E}_{t_{-i}|\beta^0_{-i}}\left\{u_i\left(\omega\left(t_i, t_{-i}\right), t\right)\right\} \geq \mathbb{E}_{t_{-i}|\beta^0_{-i}}\left\{u_i\left(\omega\left(t_i', t_{-i}\right), t\right)\right\}. \tag{3.5}$$

Accordingly, if $\omega$ is incentive compatible, it also satisfies the interim incentive compatibility condition in (3.5). However incentive compatibility is generally a more demanding condition than (3.5) since it requires truthful reporting to be optimal not only at the interim stage (under the prior $\beta^0_{-i}$), but also at the post- revelation stage (under all equilibrium path posteriors $\beta_{-i}$ in the support of $\Pi_{-i}$).[15]

[15]We will shortly demonstrate that individual rationality and incentive compatibility characterize the equilibrium outcome functions of the reciprocal contracting game. One class of outcome functions which are both individually rational and incentive compatible are the ones jointly satisfying conditions (3.2) and (3.5). Any such outcome function is individually rational and incentive compatible under the posterior system that puts unit mass on the prior $\beta^0_{-i}$. These outcomes can be supported by equilibria where no information is revealed with the equilibrium contract offers. The earlier literature on mechanism design is based on the premise that these are the only outcomes to be expected when players get together to negotiate a default game. However, as we have seen

3.2.1. ***Incentive Compatibility and the Cournot Game.*** We turn to the Cournot example one more time to demonstrate how one can verify 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 cost functions are linear, this maximization requires that, given any type profile, the player with the higher cost produces zero output and the other player produces his *monopoly output*. The resulting output levels are as in the table below:

| | $t_2 = 0.65$ | $t_2 = 1$ |
|---|---|---|
| $t_1 = 0.6$ | $y_1^* = 0.2, y_2^* = 0$ | $y_1^* = 0.2, y_2^* = 0$ |
| $t_1 = 0.7$ | $y_1^* = 0, y_2^* = 0.175$ | $y_1^* = 0.15, y_2^* = 0$ |

Under the posterior system $\Pi^*$, which we defined above, the expected non-cooperative payoff of player 1 is $93/3200$ for type 0.6 and $47/3200$ for type 0.7. Suppose the outcome function induces transfers between the players such that the resulting payoffs for the two types of player 1 equal to these lower bounds, regardless of the type of player 2. Let $\Delta z^*$ be $z_1 - z_2$[16]

| | $t_2 = 0.65$ | $t_2 = 1$ |
|---|---|---|
| $t_1 = 0.6$ | $\Delta z^* = 35/3200$ | $\Delta z^* = 35/3200$ |
| $t_1 = 0.7$ | $\Delta z^* = -47/3200$ | $\Delta z^* = 25/3200$ |

We label the resulting outcome function as $\omega^*$. By construction, $\omega^*$ satisfies the individual rationality constraints of player 1 under $\Pi^*$. Player 2's expected payoff from $\omega^*$ is $\frac{1}{2}(35/3200) + \frac{1}{2}(25/3200) = 30/3200$ for type 1 and $\frac{1}{2}(35/3200) + \frac{1}{2}(51/3200) = 43/3200$ for type 0.65.[17] These figures satisfy the individual rationality constraints for player 2 under $\Pi^*$ given in Section 3.1.1, which does not involve any first round revelation by player 1. We conclude that $\omega^*$ is individually rational. Observe that

in Section 3.1, there are outcome functions which violate condition (3.2) and yet which are still classified as individually rational under some non-degenerate posterior system.

[16]The monopoly profit is $(0.4)^2 = 128/3200$ under cost 0.6 and $(0.15)^2 = 72/3200$ under cost 0.7. The transfers on the table below sets the type dependent payoff of player 1 at the targeted level. We only report the *net* transfers the gross values of $z_1$ and $z_2$ are redundant.

[17]The monopoly profit is $(0.175)^2 = 98/3200$ under cost 0.65. This yields the payoff 43/3200 net of the transfer when player 2 type 0.65 faces player 1 type 0.7.

the outcome function $\omega^*$ is interim incentive compatible as well, since it satisfies (3.5) for both players.[18]

Now consider function $\omega^*_\Pi$, which is the natural extension of the outcome function $\omega^*$ under posterior system $\Pi^*$. Notice that, under $\Pi^*$, extended outcome function $\omega^*_\Pi$ satisfies the pre revelation incentive compatibility constraints as equalities, since $\omega^*_\Pi$ is constant in $\beta$. Moreover, since $\Pi^*$ does not prescribe any first round information revelation by player 1, the post revelation incentive compatibility constraints for player 2 are identical to interim incentive compatibility constraints. These constraints are satisfied by $\omega^*_\Pi$ since $\omega^*$ is interim incentive compatible. Yet $\omega^*_\Pi$ does not satisfy this condition for player 1. To see this, suppose player 1 updated his belief to $\beta_2 = 0$ after the first round of the game. That is, player 2 revealed his type as the high cost type with his revelation. In this case, if player 1 with type 0.6 imitates type 0.7, he receives the continuation payoff

$$47/3200 + y_1^*(0.7, 1)(0.7 - 0.6) = 95/3200$$

which is higher than the payoff from the outcome function 93/3200.

The example above indicates that post revelation incentive compatibility is generally stronger than interim incentive compatibility. However, we can construct an

[18] This follows from the four inequalities below:

$$\begin{aligned}
u_1\left(E_{t_2|\beta_2^0}\omega^*(0.6,t_2),0.6\right) - u_1\left(E_{t_2|\beta_2^0}\omega^*(0.6,t_2),0.7\right) &\geq u_1\left(E_{t_2|\beta_2^0}\omega^*(0.6,t_2),0.6\right) - u_1\left(E_{t_2|\beta_2^0}\omega^*(0.7,t_2),0.7\right)\\
u_1\left(E_{t_2|\beta_2^0}\omega^*(0.6,t_2),0.6\right) - u_1\left(E_{t_2|\beta_2^0}\omega^*(0.7,t_2),0.7\right) &\geq u_1\left(E_{t_2|\beta_2^0}\omega^*(0.7,t_2),0.6\right) - u_1\left(E_{t_2|\beta_2^0}\omega^*(0.7,t_2),0.7\right)\\
u_2\left(E_{t_1|\beta_1^0}\omega^*(t_1,0.65),0.65\right) - u_2\left(E_{t_1|\beta_1^0}\omega^*(t_1,0.65),1\right) &\geq u_2\left(E_{t_1|\beta_1^0}\omega^*(t_1,0.65),0.65\right) - u_2\left(E_{t_1|\beta_1^0}\omega^*(t_1,1),1\right)\\
u_2\left(E_{t_1|\beta_1^0}\omega^*(t_1,0.65),0.65\right) - u_2\left(E_{t_1|\beta_1^0}\omega^*(t_1,1),1\right) &\geq u_2\left(E_{t_1|\beta_1^0}\omega^*(t_1,1),0.65\right) - u_2\left(E_{t_1|\beta_1^0}\omega^*(t_1,1),1\right)
\end{aligned}$$

The first two inequalities simplify to

$$\begin{aligned}
(0.7-0.6)\,E_{t_2|\beta_2^0}y_1^*(0.6,t_2) &\geq 46/3200 \geq (0.7-0.6)\,E_{t_2|\beta_2^0}y_1^*(0.7,t_2)\\
(0.1)(0.2) &\geq 46/3200 \geq (0.1)(0.075)
\end{aligned}$$

The last two inequalities simplify to

$$\begin{aligned}
(1-0.65)\,E_{t_1|\beta_1^0}y_2^*(t_1,0.65) &\geq 13/3200 \geq (1-0.65)\,E_{t_1|\beta_1^0}y_2^*(t_1,1)\\
(0.1)(0.0875) &\geq 13/3200 \geq (0.1)\,0
\end{aligned}$$

alternative extended outcome function $\tilde{\omega}_{\Pi}^{*}$ which is also consistent with the outcome function $\omega^{*}$ and which is both pre revelation and post revelation incentive compatible. As the natural extension $\omega_{\Pi}^{*}$, function $\tilde{\omega}_{\Pi}^{*}$ also induces the joint profit maximizing levels of the output $y_1^*$ and $y_2^*$ for each type profile. However, unlike the monetary transfers in $\omega_{\Pi}^{*}$, the transfers in $\tilde{\omega}_{\Pi}^{*}$ depend on the first round revelation of player 2. Let $\Delta\tilde{z}^{*}$ be the difference between the $z_1$ and $z_2$ in $\tilde{\omega}_{\Pi}^{*}$.

$$\Delta\tilde{z}^{*}\left(t_1, t_2, \beta_2\right) = \Delta z^{*}\left(t_1, t_2\right) + x\left(t_1, \beta_2\right)$$

where $x\left(t_1, \beta_2\right)$ is defined as follows

$$x\left(0.6, \beta_2\right) = -x\left(0.7, \beta_2\right) = \begin{cases} 5.4/3200 \text{ if } \beta_2 = 0 \\ -3.24/3200 \text{ if } \beta_2 = 0.8 \end{cases}$$

Let us summarize our discussion on individual rationality and incentive compatibility. The possibility of information revelation through contracts makes some outcome functions satisfy individual rationality even though they yield a payoff lower than what a player would expect from the default game under his prior beliefs. However, information revelations at the contracting stage come at the cost of stronger incentive compatibility requirements than the standard interim conditions.

Accommodating for information revelation through contract offers complicates the definitions of the IR and IC constraints. Stating these in our setting requires referring to systems of posteriors and extended outcome functions, which do not appear in the standard interim versions of these conditions. However, this does not mean that we could get a simpler characterization 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 (forthcoming) show that even a simple yes or no decision on a third party's mechanism can reveal the type of the responding player and therefore extend the set of feasible allocation rules beyond what is outlined by the standard interim IR and IC conditions.

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

**Theorem 1.** *An outcome function $\omega$ is supportable as a perfect Bayesian equilibrium outcome function of the reciprocal contracting game if and only if there exist refusal beliefs $\beta^{no}$ and a posterior system $\Pi$ under which $\omega$ is individually rational and incentive compatible.*

We prove Theorem 1 together with Proposition 1. 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 $\beta^{no}$ and some posterior system $\Pi$. This step proves the only if direction of the theorem. For the second part, we start with an outcome function $\omega$ which is individually rational and incentive compatible under some $\beta^{no}$ and $\Pi$. We construct an equilibrium which supports the outcome function $\omega$ and which satisfies conditions (i) to (iv) of Proposition 1, proving the proposition and the if direction of the theorem.

*Proof.* **PART I:**

Suppose there exists a perfect Bayesian equilibrium of the reciprocal contracting game such that $\omega$ is the equilibrium outcome function. First, we will construct $\Pi$, $\beta^{no}$, and $\omega_\Pi$, by using the properties of the equilibrium. Then we will show that $\omega$ satisfies the individual rationality and incentive compatibility constraints together with the constructed $\Pi$, $\beta^{no}$, and $\omega_\Pi$.

**1) Construction of $\Pi$, $\beta^{no}$, and $\omega_\Pi$**

Consider the equilibrium path contract offers of 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 $\Pi_i$. Since the ex ante expectation over the posteriors equals the prior beliefs, distribution $\Pi_i$ is Bayes plausible. The system of posteriors $\Pi$ is defined as $\{\Pi_i\}_{i\in I}$.

There are infinitely many possible mechanisms for each player and therefore there are infinitely many $\delta$ 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 $\beta_i^{no}$ be the (possibly off the equilibrium path) posterior belief on player $i$'s type following the observation of this contract. $\beta^{no}$ 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 $\beta$ which is in the support of $\Pi$. 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 $\omega_\Pi(\cdot, \beta)$ be this mapping. Bayes rule implies that $\omega_\Pi$ induces $\omega$.

**2) Verifying the constraints**

We now need to show that $\omega$, $\Pi$, $\beta^{no}$, and $\omega_\Pi$ together satisfy the individual rationality and incentive compatibility constraints.

The right hand side of the $IR$ constraint 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 $\beta_{-i}$ (which depend on the other players' contracts) and $\beta_i^{no}$. 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 $IR$ constraint must hold.

It follows from the construction of the posterior system $\Pi$ that any distribution $\beta_i$ in the support of $\Pi_i$ 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 $IC$ constraint corresponds to the expected payoff from offering the contract corresponding to posterior $\beta_i' \in supp(\Pi_i)$ and then continuing with the equilibrium continuation play. For this strategy not to be a profitable deviation, pre-revelation $IC$ must be satisfied.

Similarly, post-revelation $IC$ 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 $\omega$ which satisfies the individual rationality and incentive compatibility conditions together with some $\Pi$ and $\beta^{no}$. Incentive compatibility implies existence of an extended outcome function $\omega_\Pi$ which is consistent with $\omega$ and which satisfies conditions (3.3) and (3.4). By using these objects, we will first construct a profile of strategies and beliefs satisfying conditions (i) - (iv) defined in Proposition 1. Then we will argue that these strategies induce $\omega$ as the outcome function. Finally, we will show that the strategies and beliefs we constructed constitute a Perfect Bayesian Equilibrium of the reciprocal contracting game.

**1) Strategies and Beliefs**

**a) Equilibrium contracts:**

Recall that a contract by a player $i$ consists of a 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 $\Pi_i$. Since $\Pi$ is a belief system, $\Pi_i$ is a Bayes plausible distribution of posteriors in 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 $\Pi_i$. Yet, in order to fully define function $\delta^*$, 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 $\Pi_i$, the equilibrium contracts interpret this as if this player made an arbitrary revelation within the support of $\Pi_i$. To formalize this idea, we let $\beta_i^1$ be an arbitrary posterior in the support of the distribution $\Pi_i$ and define a transformation function $\bar{\beta}_i : \Delta T_i \rightarrow supp\,(\Pi_i)$ 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 $\left\{\bar{\beta}_i\left(\hat{\beta}_i\right)\right\}_{i\in I} \in supp\,(\Pi)$.

Index the action profiles in $A$ in some arbitrary way. Let $\omega_\Pi^k(t,\beta)$ be the probability assigned to action profile $a^k$ by the extended outcome function $\omega_\Pi$ when player types are $t$ and the posterior is $\beta$. The notation $a_i^k$ stands for the action taken by player $i$ in action profile $a^k$.

We are now ready to state the list of mechanisms players will submit as part of their reciprocal contracts. Recall that a mechanism for player $i$ maps the type reports $t \in T$ and correlating messages $x = \{x_i\}_{i\in I} \in [0,1]^{|I|}$ to an action in $A_i$. When revelations by the players is $\hat{\beta} \in \times_{i\in I}\Delta T_i$, function $\delta^*$ determines the mechanisms in the second round according to the following formula:

$$m_i^{\hat{\beta}}(t,x) = \left\{a_i^k : k = \min_{k'} \text{ such that } \sum_{\tau=1}^{k'} \omega_\Pi^k\left(t, \bar{\beta}\left(\hat{\beta}\right)\right) \geq \left\lfloor \sum_j x_j \right\rfloor \right\}. \tag{3.6}$$

In this formula, $\left\lfloor \sum_j x_j \right\rfloor$ refers to the fractional part of the real number $\sum_j x_j$ or to $\sum_j x_j$ in mod 1. This function aggregates the numbers sent by the players into another number in the unit interval and uses this number to choose the index of the action profile in $A$, depending on the type reports and the posterior distribution. The resulting mechanism directs player $i$ to take his part in this action profile.

**b) Equilibrium beliefs:**

After the first round of the game, all players observe the contract offers. Description of an equilibrium demands for 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 $\delta^*$ described above and a revelation $\hat{\beta}_i$ in the support of $\Pi_i$. After observing this offer, abiding by the Bayes rule, the other players update their belief on this players 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 $\delta^*$ and a revelation $\hat{\beta}_i \notin \Pi_i$; and to the refusal belief $\beta_i^{no}$ when player $i$'s contract consists of a list other than $\delta^*$.

then the beliefs on $i$ are updated to $\beta_i^1 \in supp(\Pi_i)$.

If contract of player $i$ includes the list $\delta^*$, then the other players update their belief on this player's type to $\bar{\beta}_i\left(\hat{\beta}_i\right)$. In words, if this player submits a revelation $\hat{\beta}_i$ in the support of $\Pi_i$, the other players update their belief to $\hat{\beta}_i$ assuming that he is "truthful." Otherwise, when he submits a revelation outside the support of $\Pi_i$, the other players change their belief to $\beta_i^1 \in supp(\Pi_i)$. On the other hand, if the contract of player $i$ includes a list of mechanisms other than $\delta^*$, then the belief on this player's type is updated to the refusal belief $\beta_i^{no}$.

**c) Equilibrium reports to direct mechanisms which are induced by $\delta^*$:**

Suppose, in the first round, all players offer contracts including the list $\delta^*$ we described above. In the second round, each player $i$ should submit a type report $t_i$ and a correlating number $x_i$ to the resulting mechanisms. In the equilibrium we construct, $t_i$ equals the type of player $i$ and $x_i$ is uniformly distributed on the interval $[0, 1]$, 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 $\delta^*$ described above. We set these as arbitrary.[19]

**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 $\omega$. We start with the equilibrium path subgame that begins when players all submit the list $\delta^*$ and their revelations are $\hat{\beta} \in supp(\Pi)$.

[19]Reaching to these decision nodes requires all players to deviate from equilibrium behavior and our solution concept does not impose any requirement on them.

The equilibrium strategies prescribe that each player reports his type truthfully and sends a correlating message which is generated by the uniform distribution on the interval $[0,1]$. What is crucial in our derivation is that the function $\lfloor \cdot \rfloor$ maps the sum of uniformly distributed random variables on $[0,1]$ to another uniformly distributed random variable on the same interval. Formally, $\lfloor \sum_{i \in I} x_i \rfloor$ is uniform on $[0,1]$ as long as $x_i$ is uniform on $[0,1]$ for all $i$. Therefore, when players report their true types to the mechanisms in (3.6), these mechanisms together implement action profile $a^k$ with probability $\omega_\Pi^k (t, \beta)$ as the extended outcome function requires. Now recall that each player $i$ determines his first round revelation message $\hat{\beta}_i$ in a way to support $\Pi_i$ as the distribution over the posteriors on his type. The proof follows from the fact that the extended outcome function $\omega_\Pi$ is consistent with $\omega$:

$$\omega(t) = \mathbb{E}_{\hat{\beta}|\Pi} \omega_\Pi \left( t, \hat{\beta} \right).$$

**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.

The next step is showing that there is no profitable deviation from the behavior described in (c) given the beliefs in (b). Let us start with the players' choice of their correlating messages. If player $i$ expects that all the other players choose their correlating messages with respect to the uniform distribution, then $\left\lfloor \sum_{j \neq i} x_j \right\rfloor$ is uniformly distributed on $[0,1]$. Moreover, for all $x_i \in [0,1]$, the random variable $\left\lfloor x_i + \sum_{j \neq i} x_j \right\rfloor$ has a uniform distribution on the same interval as well. This proves that player $i$ is indifferent between all the correlating messages in his disposal. This means that whatever else is happening, it is a best reply for each player $i$ to select a number $x_i$ using a uniform distribution provided he believes the others are doing the same thing. The optimality of revealing the true types under beliefs $\bar{\beta} \left( \hat{\beta} \right)$ follows from the post-revelation incentive compatibility (3.4) of the extended outcome function $\omega_\Pi$.

As a final step to our proof, we consider the 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 $\delta^*$ together with some revelation $\hat{\beta}_i'$ such that $\hat{\beta}_i' \in supp\ (\Pi_i)$ and $\hat{\beta}_i'(t_i) = 0$.[20] Pre-revelation incentive compatibility (3.3) of the extended allocation function $\omega_\Pi$ implies that this is not a profitable deviation. Secondly, a player $i$ with type $t_i$ may choose to offer a contract which includes a list other than $\delta^*$. According to the beliefs in (b), all players change their belief on player $i$ to $\beta_i^{no}$, and the beliefs on the other players are determined by the posterior system $\Pi$. After this deviation, all players follow their non-cooperative mechanisms in the second round. For player $i$, this continuation behavior yields an expected payoff equal to the right hand side of the constraint (3.1). Individual rationality of the outcome function $\omega$ implies that this is not a profitable deviation either. $\square$

## 4. Incentives under Private Values and Single Crossing

The characterization result of the previous section suggests the following procedure to confirm that an outcome function $\omega$ is supportable with a Perfect Bayesian equilibrium. First, find the refusal beliefs and posterior systems under which $\omega$ is individually rational. Then, investigate if, for any of these posterior systems, one can construct an extended outcome function which is consistent with $\omega$ and which satisfies the pre revelation and post revelation incentive compatibility conditions. This procedure is not very practical for many settings. Fortunately, for an important class of design environments, we can simplify the verification of the incentive compatibility requirements. Specifically, under the assumptions of private values and single crossing, we can show that our incentive compatibility conditions essentially boil down to the less demanding and more familiar interim incentive compatibility constraints.

A player'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.

[20] A revelation outside of $supp\ (\Pi_i)$ is strategically equivalent to the revelation $\beta_i^1$ which is in $supp\ (\Pi_i)$.

To describe the second condition we impose, we relabel types of player $i$ such that $T_i \subset \mathbb{R}$. We let $h_i$ be a function from randomizations over action profiles into real numbers, i.e, $h_i : \Delta A \to \mathbb{R}$. Preferences of player $i$ satisfy the **single crossing property** with function $h_i(\cdot)$ if

$$h_i(q) \geq h_i(q') \text{ if and only if } u_i(q, t_i) - u_i(q', t_i) \text{ is weakly decreasing in } t_i.$$

For many settings, where preferences satisfy a one dimensional condensation condition (Mookherjee and Reichelstein, 1992), function $h_i(\cdot)$ has a natural interpretation. For instance, in our Cournot game, players' preferences satisfy the single crossing condition where $h_i$ is the production level of player $i$. Similarly, this condition holds for private value auctions, with $t_i$ as player $i$'s valuation and $h_i$ as minus the probability that he receives the auctioned object, and for public good provision games, with $t_i$ as player $i$'s valuation and $h_i$ as minus the total amount of the public good. Single crossing property allows the players to use the level of $h_i$ to differentiate between different types.

Under the single crossing property, incentive compatibility demands a monotonic relationship between $t_i$ and $h_i$. If player $i$'s preferences exhibit private values and satisfy single crossing with function $h_i(\cdot)$, then the interim incentive compatibility condition (3.5) implies that $h_i\left[\mathbb{E}_{t_{-i}|\beta^0_{-i}}\omega(t_i, t_{-i})\right]$ is weakly decreasing in $t_i$. Moreover, many of the incentive compatibility constraints are redundant under these conditions: If the interim (or post revelation) incentive compatibility constraints are satisfied between all the "adjacent" types of player $i$, then all the other interim (or post revelation) incentive compatibility constraints (between the non-adjacent types of player $i$) hold as well.

Now we describe a structure which can be used by the players to transfer utility to each other. Given a posterior system $\Pi$, we define function $x_i(t, \beta)$ as the transfer to player $i$ when the realized type profile and posterior belief are $t$ and $\beta$ 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 $t, \beta$; and **outcome neutral**, i.e, $\mathbb{E}_{\beta|\Pi} x_i(t, \beta) = 0$. Under the system of posteriors $\Pi$, an extended outcome function $\omega_\Pi$ is **pre revelation incentive compatible with transfer rule** $x$ if

$$
\begin{aligned}
& \mathbb{E}_{\beta_{-i}|\Pi_{-i}} \mathbb{E}_{t_{-i}|\beta_{-i}} \left\{ u_i \left( \omega_\Pi \left( t, \beta_i, \beta_{-i} \right), t \right) + x_i \left( t, \beta_i, \beta_{-i} \right) \right\} \\
\text{(4.1)} \qquad \geq \; & \mathbb{E}_{\beta_{-i}|\Pi_{-i}} \mathbb{E}_{t_{-i}|\beta_{-i}} \left\{ u_i \left( \omega_\Pi \left( t, \beta_i', \beta_{-i} \right), t \right) + x_i \left( t, \beta_i', \beta_{-i} \right) \right\}
\end{aligned}
$$

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 system of posteriors $\Pi$, an extended outcome function $\omega_\Pi$ is **post revelation incentive compatible with transfer rule** $x$ if for all $\beta \in supp(\Pi)$,

(4.2)

$$
\mathbb{E}_{t_{-i}|\beta_{-i}} \left\{ u_i \left( \omega_\Pi \left( t_i, t_{-i}, \beta \right), t \right) + x_i \left( t, \beta_i, \beta_{-i} \right) \right\} \geq \mathbb{E}_{t_{-i}|\beta_{-i}} \left\{ u_i \left( \omega_\Pi \left( t_i', t_{-i}, \beta \right), t \right) + x_i \left( t, \beta_i, \beta_{-i} \right) \right\}
$$

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

In the context of the Cournot game, we have already seen an example to how a transfer rule may help fulfilling the incentive compatibility requirements of reciprocal contracting. In the example, we started with the interim incentive compatible allocation function $\omega^*$. This function was also individually rational under the refusal beliefs $\beta^{no*}$ and the posterior system $\Pi^*$. However, its natural extension under $\Pi^*$ did not satisfy one of the post revelation incentive compatibility constraints. Nevertheless, we showed the existence of monetary transfers between the players, which would support the allocation function $\omega^*$ and satisfy the pre revelation and post revelation incentive compatibility conditions as well. Our next result extends this observation to all Bayes plausible posterior systems and all settings satisfying conditions of private values and single crossing.

**Theorem 2.** *Suppose all the players' preferences exhibit private values and satisfy the single crossing condition with functions $\{h_i(\cdot)\}_{i \in I}$. Suppose further that $\omega$ is an interim incentive compatible outcome function with the property that $h_i(\omega(t_i, t_{-i}))$ is weakly decreasing in $t_i$ for all $t_{-i}$ and all $i$. Let $\Pi$ be an arbitrary posterior system and $\omega_\Pi$ be the natural extension of $\omega$ given this system. There exists a transfer rule $x$ such that $\omega_\Pi$ is pre revelation and post revelation incentive compatible with $x$ under posterior system $\Pi$.*

*Proof.* Let $t_i$ and $t_i'$ be two adjacent types of player $i$. Suppose player $i$ has posterior $\beta_{-i}$ on the types of the other players. Under this belief, define $\Delta_i^{\beta_{-i}}(t_i, t_i')$ as the utility premium of type $t_i$ from revealing his type truthfully instead of imitating type $t_i'$:

$$\Delta_i^{\beta_{-i}}(t_i,t_i') = \mathbb{E}_{t_{-i}|\beta_{-i}}\left[u_i\left(\omega\left(t_i,t_{-i}\right),t_i\right) - u_i\left(\omega\left(t_i',t_{-i}\right),t_i\right)\right].$$

The next step is defining function $g_i\left(t_i,\beta_{-i}\right)$. For fixed $\beta_{-i}$, the rate of change of function $g_i$ between any two adjacent types $t_i$ and $t_i'$ is given as:

$$g_i\left(t_i,\beta_{-i}\right) - g_i\left(t_i',\beta_{-i}\right) = \frac{\Delta_i^{\beta_{-i}}(t_i',t_i)\,\Delta_i^{\beta_{-i}^0}(t_i,t_i') - \Delta_i^{\beta_{-i}}(t_i,t_i')\,\Delta_i^{\beta_{-i}^0}(t_i',t_i)}{\Delta_i^{\beta_{-i}^0}(t_i,t_i') + \Delta_i^{\beta_{-i}^0}(t_i',t_i)}. \tag{4.3}$$

This equation determines $g_i\left(\cdot,\beta_{-i}\right)$ up to a constant. The following equation yields this constant term and thus pins down the function:

$$\mathbb{E}_{t_i|\beta_i^0} g_i\left(t_i,\beta_{-i}\right) = 0. \tag{4.4}$$

Bayes plausibility of posterior system $\Pi$ implies that, $\Delta_i^{\beta_{-i}}(t_i,t_i')$ equals $\Delta_i^{\beta_{-i}^0}(t_i',t_i)$ in expectation given $\Pi$:

$$\mathbb{E}_{\beta_{-i}|\Pi_{-i}}\Delta_i^{\beta_{-i}}(t_i,t_i') = \Delta_i^{\beta_{-i}^0}(t_i',t_i) \text{ if } \mathbb{E}_{\beta_{-i}|\Pi_{-i}}\beta_{-i} = \beta_{-i}^0.$$

In light of this property, take the expectation of both sides of equations (4.3) and (4.4) over the variable $\beta_{-i}$ given a Bayes plausible posterior system $\Pi_{-i}$ :

$$\begin{aligned} \mathbb{E}_{\beta_{-i}|\Pi_{-i}} g_i\left(t_i,\beta_{-i}\right) - \mathbb{E}_{\beta_{-i}|\Pi_{-i}} g_i\left(t_i',\beta_{-i}\right) &= 0 \\ \mathbb{E}_{t_i|\beta_i^0}\mathbb{E}_{\beta_{-i}|\Pi_{-i}} g_i\left(t_i,\beta_{-i}\right) &= 0 \end{aligned}$$

These equations show that $\mathbb{E}_{\beta_{-i}|\Pi_{-i}} g_i\left(t_i,\beta_{-i}\right) = 0$ for all $t_i$.

We define the transfer $x_i$ as

$$x_i\left(t_i,t_{-i},\beta_i,\beta_{-i}\right) = g_i\left(t_i,\beta_{-i}\right) - \frac{1}{|I|-1}\sum_{j\neq i} g_j\left(t_j,\beta_{-j}\right)$$

Notice that $\sum_i x_i = 0$. Moreover $\mathbb{E}_{\beta_{-i}|\Pi_{-i}} g_i\left(t_i,\beta_{-i}\right) = 0$ implies $\mathbb{E}_{\beta|\Pi} x_i(t,\beta) = 0$ for any posterior system $\Pi$. This means that $x = \{x_i\}$ is a well defined transfer rule.

Define the extended outcome function as $\omega_\Pi(t,\beta) = \omega(t)$ for all $t,\beta$.

Pre-revelation IC constraints are satisfied as equalities:

$$\mathbb{E}_{\beta_{-i}|\Pi_{-i}}\mathbb{E}_{t_{-i}|\beta_{-i}} x_i\left(t_i,t_{-i},\beta_i,\beta_{-i}\right) = \mathbb{E}_{\beta_{-i}|\Pi_{-i}} g_i\left(t_i,\beta_{-i}\right) - \frac{1}{|I|-1}\sum_{j\neq i}\mathbb{E}_{\beta_{-i-j}|\Pi_{-i-j}}\mathbb{E}_{t_j|\beta_j^0} g_j\left(t_j,\beta_{-j}\right) = 0$$

since $\mathbb{E}_{\beta_{-i}|\Pi_{-i}} g_i\left(t_i, \beta_{-i}\right) = 0$ for all $t_i$ and $\mathbb{E}_{t_j|\beta_j^0} g_j\left(t_j, \beta_{-j}\right) = 0$ for all $\beta_{-j}$.

Post-revelation IC constraints take the form:

$$\Delta_i^{\beta_{-i}}\left(t_i, t_i'\right) + g_i\left(t_i, \beta_{-i}\right) - g_i\left(t_i', \beta_{-i}\right) \geq 0$$

for all $\beta_{-i}$. Using equation (4.3), the left hand side equals:

$$\frac{\Delta_i^{\beta_{-i}}\left(t_i, t_i'\right) + \Delta_i^{\beta_{-i}}\left(t_i', t_i\right)}{\Delta_i^{\beta_{-i}^0}\left(t_i, t_i'\right) + \Delta_i^{\beta_{-i}^0}\left(t_i', t_i\right)} \Delta_i^{\beta_{-i}^0}\left(t_i, t_i'\right)$$

Interim IC implies $\Delta_i^{\beta_{-i}^0}\left(t_i, t_i'\right)$ and $\Delta_i^{\beta_{-i}^0}\left(t_i', t_i\right)$ are non-negative. Notice that

$$\Delta_i^{\beta_{-i}}\left(t_i, t_i'\right) + \Delta_i^{\beta_{-i}}\left(t_i', t_i\right) = \mathbb{E}_{t_{-i}|\beta_{-i}} \left\{ \begin{array}{c} \left[u_i\left(\omega\left(t_i, t_{-i}\right), t_i\right) - u_i\left(\omega\left(t_i', t_{-i}\right), t_i\right)\right] \\ -\left[u_i\left(\omega\left(t_i, t_{-i}\right), t_i'\right) - u_i\left(\omega\left(t_i', t_{-i}\right), t_i'\right)\right] \end{array} \right\}$$

The term inside the set brackets is larger than zero for all $t_{-i}$, since $h_i\left(\omega\left(t_i, t_{-i}\right)\right)$ is weakly decreasing. $\square$

## References


[1] Celik G. and M. Peters, forthcoming. "Equilibrium rejection of a mechanism" Games and Economic Behavior.

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

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

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

[5] Farell, J., and C. Shapiro, 1990. "Horizontal Mergers: An Equilibrium Analysis" American Economic Review, 80(1), 107–126.

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

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

[8] Mookherjee, D. and S. Reichelstein, 1992. "Dominant strategy implementation of Bayesian incentive compatible allocation rules" Journal of Economic Theory, 56, 378 - 399.

[9] Myerson, R., 1983. "Mechanism Design by an Informed Principal," Econometrica, 51, 1767–1797.

[10] Peters, M., 2010. "On the Revelation Principle and Reciprocal Mechanisms in Competing Mechanism Games," Discussion paper, University of British Columbia.

[11] Peters, M., and B. Szentes, 2008 "Definable and Contractible Contracts," Working paper, University of British Columbia.

[12] 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.