

# IMPLEMENTATION OF OPTIMAL CONTRACTS UNDER ADVERSE SELECTION\*

**Pablo Amorós and Bernardo Moreno\*\***

WP-AD 98-25

Correspondence to:

Pablo Amorós. University of Alicante. Facultad de Ciencias Económicas.

Dpto. Fundamentos del Análisis Económico. Ctra. San Vicente del Raspeig, s/n.

03071 ALICANTE-SPAIN

E-mail: pamoros@merlin.fae.ua.es

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.

First Edition December 1998

ISBN: 84-482-1973-2

Depósito Legal: V-4883-1998

IVIE working-papers offer in advance the results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication.

---

\* We thank J. López, I. Macho, I. Ortuño-Ortín and two anonymous referees for their comments. The discussions with our thesis adviser L. Corchón were specially useful.

\*\* P. Amorós and B. Moreno: University of Alicante.

# IMPLEMENTATION OF OPTIMAL CONTRACTS UNDER ADVERSE SELECTION

**Pablo Amorós and Bernardo Moreno**

## **Abstract**

This paper studies the principal-multiagent model of a firm subject to adverse selection. We focus on agents who have complete information. We propose some desirable properties to be satisfied by mechanisms implementing the first-best in Nash equilibrium: (i) enforceability (a property related with the individual rationality of the mechanism), (ii) renegotiation-proofness, (iii) small strategy spaces, (iv) unique implementation, (v) unique best-reply, and (vi) no mixed strategies. We prove that enforceability is not compatible with renegotiation-proofness or the unique best-reply property. Then we propose two mechanisms. The first one satisfies properties (i), (ii) and (iv). The second mechanism verifies all properties but enforceability.

**Keywords:** Adverse Selection, Contract Theory, Implementation Theory, Nash Equilibrium.

**JEL Classification Numbers:** C720, D820

## 1. INTRODUCTION

The principal-agent model has been extensively studied in the literature (see Macho and Pérez Castrillo [10] for a good review). Specifically, the principal-two agents problem of a firm subject to adverse selection has been studied in several papers (Demski and Sappington [2], Ma, Moore and Turnbull [9] and Glover [6]). These authors assume that agents have private information about their types and that agents' types are imperfectly correlated. The first-best contracts are those that the principal would like to sign if he knew agents' types. Since these contracts are such that some type of agents have incentives to lie, the Revelation Principle (Gibbard [5]) implies that they are not implementable either in dominant strategies or in Bayes-Nash equilibrium. Then, most papers focus on designing mechanisms that implement the second-best contracts (those that maximize the expected profits of the principal subject to truth-telling is an equilibrium) in Bayes-Nash equilibrium.

In this paper, we assume that agents have complete information about themselves and their colleagues (think of workers of a section in some firm, or of professors of a department in a University). We show that the principal can get advantage from this informational structure among agents in the sense that he can implement the first-best contracts in Nash equilibrium (Demski and Sappington [3] consider the case of perfect correlation of types, a special case of complete information).

If we want to be sure that the first-best contracts are decentralizable, it is not enough to provide a mechanism that implements them in the formal sense (i.e. for each state of the world, the set of Nash equilibria outcomes coincides with the first-best contracts). Think for example on the mechanisms used in the proofs of general results establishing necessary and sufficient conditions for implementation in Nash equilibrium (see Maskin [11], Moore and Repullo [13], and Dutta and Sen[4]). They implement in Nash equilibrium any choice rule satisfying these conditions. However, the complexity of that mechanisms makes it hard to believe that they can be applied in the real world.

In order to put in practice a mechanism it should verify some properties that guarantee that agents will play it just as we expect. From here, we have consider two different factors. First, the mechanism should take into account the way in which agents calculate their expected utility by playing some profile of strategies. This is fundamental to know which are the real equilibria of the game. Secondly, we have to be sure that agents will be able to find the equilibria of the mechanism

without many problems.

Keeping in mind these two factors, we propose six desirable properties to be satisfied by any mechanism implementing the first-best contracts in Nash equilibrium: enforceability, renegotiation-proofness, small strategy spaces, unique implementation, unique best-reply property and no mixed strategies.

The first two properties refer to the way in which agents calculate their utility by playing each profile of strategies. Enforceability is a property related with the individual rationality of the mechanism. The implementation literature has paid little attention to this problem. Suppose that for some profile of strategies the mechanism selects outcomes that are not individually rational for any type of agent (i.e. outcomes that give to the agents a utility level smaller than the reservation utility level whatever their types are). In that case agents would prefer to reject these outcomes. Obviously, if they can do that, their expected utility by playing these strategies is not given by the outcomes that the mechanism selects. To avoid that situation we will demand that the mechanisms always choose contracts that can not be rejected by agents. Of course, this means a restriction on the outcomes that can be selected by the mechanisms<sup>1</sup>.

Suppose now that, once the mechanism has selected some outcomes, the agents renegotiate among them. Again, if these renegotiation takes place, the agents' expected utility by playing any profile of strategies is not given by the outcomes that the mechanism chooses, but it will be given by the outcomes that they obtain after the renegotiation. The implementing mechanisms should take this into account. Usually, in the implementation literature studying this type of problems, the renegotiation rule is taken as given (see Moore [12]). However, in general we may have no idea about which is this renegotiations rule. A mechanism is renegotiation-proof when it works whatever the renegotiation rule is.

The objective of the other four properties is to make the search for the equilibrium strategies easier. For that, we first require that the strategy spaces be as small as possible. Like this, the rules of the mechanism are more simple, and agents have less problems to coordinate among them to play the Nash equilibrium strategies. In the same spirit, the unique implementation property claims that, for each state of the world, there should not be more than one Nash equilibrium. In this way we avoid problems of coordination on which equilibrium to play. The aim of the unique-best reply property is to give to agents strong incentives to stick to the equilibrium strategies, by requiring that the equilibrium strategy of

---

<sup>1</sup>In a bargaining model, Jackson and Palfrey [8] examine this form of implementation with individually rational mechanisms (they call it voluntary implementation).

each agent must be the unique best response to the equilibrium strategies of the other agents. Finally, we also want to avoid the existence of mixed equilibrium strategies. On the one hand, the existence of mixed and pure Nash equilibrium would contradict the unique implementation and the unique best-reply properties. On the other hand, since mixed equilibrium strategies are usually hard to find, if there is only one Nash equilibrium and it involves playing mixed strategies, it is probable that the agents do not play that equilibrium (and then the mechanism would not work).

Unfortunately some of these properties are not compatible. Specifically, enforceability is a big restriction, since it reduces the set of outcomes that the mechanisms can select. We prove that this property is not compatible with renegotiation-proofness or with the unique best-reply property. Moreover, although we have not prove it, we have the intuition that enforceability is not compatible with no mixed strategies either.

Then, we propose two different mechanisms implementing the first-best contracts in Nash equilibrium. The first mechanism (Mechanism 1) is enforceable, has small strategy spaces and satisfies the unique implementation property. The second mechanism (Mechanism 2) drop out enforceability. In exchange for that it is renegotiation-proofness, has small strategy spaces, and satisfies the unique implementation, unique best-reply and no mixed strategies properties. Moreover, following the spirit of renegotiation-proofness, it satisfies the last three properties regardless of the renegotiation rule used by agents. Mechanism 2 tries to make up for the lack of enforceability by including a strategy for each agent that guarantees him the reservation utility level, regardless of the strategy announced by his counterpart.

The remainder of this paper is as follows. Section 2 presents the model and some basic definitions. In Section 3 we present the desirable properties of the mechanisms. Section 4 studies the compatibility of these properties. In Section 4 we propose the two mechanisms. Conclusions are given in Section 5.

## 2. THE MODEL

A risk-neutral principal owns two production processes,  $A$  and  $B$ . In order to operate these technologies he wants to hire two agents, agent  $A$  and agent  $B$  respectively. Each process,  $x_i = X_i(e_i, \theta_i)$ ,  $i \in \{A, B\}$ , depends on the level of effort exerted by agent  $i$ ,  $e_i$ , and a random variable with binary support,  $\theta_i \in \{\theta^1, \theta^2\}$  (with  $0 < \theta^1 < \theta^2$ ), that can be interpreted as a productivity parameter.

We will refer to this parameter as the type of the agent. When agent  $i$  is of type  $k$  ( $k \in \{1, 2\}$ ) we will write  $\theta_i^k$  (or  $\theta^k$  when it is clear enough). For all agent  $i$ , the higher  $\theta_i$  realization places the agent in a more productive setting: for all  $e_i > 0$ ,  $X_i(e_i, \theta^2) > X_i(e_i, \theta^1)$ . Therefore, when  $\theta_i = \theta^2$  ( $\theta_i = \theta^1$ ), we will say that agent  $i$  is more productive (less productive). Assume that for each process there are decreasing returns to effort ( $X_{ie} > 0$  and  $X_{iee} < 0$ , where  $X_{ie}$  and  $X_{iee}$  denotes the first and second derivative of  $X_i$  with respect to argument  $e$ ). We will also assume that for all  $e_i > 0$  and  $\theta_i \in \{\theta^1, \theta^2\}$ ,  $X_A(e_i, \theta_i) = X_B(e_i, \theta_i)^2$ .

Each agent observes  $(\theta_A, \theta_B)$  before he decides whether he signs a contract. However, the principal does not observe either  $(\theta_A, \theta_B)$  or  $(e_A, e_B)$ . Therefore contracts have to be based on the levels of outputs. A *contract* for agent  $i$  will be a pair,  $c_i = (x_i, R_i) \in \mathbb{R}_+^2$ , where  $x_i$  is a level of output produced by him, and  $R_i$  is a payment from the principal. We will assume that the price of the output is normalized to one (so  $x_i$  represents both output and value of that output). Then, when agent  $i$  signs a contract  $c_i = (x_i, R_i) \in \mathbb{R}_+^2$ , the principal's profits are  $\pi_{c_i} = x_i - R_i$ . Assume also that there is a reservation wage,  $R^0$ , denoting the payoffs that an agent would obtain if he did not sign any contract (we can think on  $R^0$  as the unemployment compensation). Notice that  $R^0$  does not depend on the type of the agent.

For all agent  $i \in \{A, B\}$  and all  $\theta_i \in \{\theta^1, \theta^2\}$ , we can compute the level of effort which is necessary for agent  $i$ ,  $e_i$ , to produce a level of output  $x_i$ . Then, agent  $i$ 's preference relation can be defined over  $x_i$  and  $R_i$ . Let  $u_i(\cdot, \theta_i) : \mathbb{R}_+^2 \rightarrow \mathbb{R}$  be a utility function representing this preference relation over the contracts in  $\mathbb{R}_+^2$ . For simplicity, we will assume that if both agents are of the same type, they are identical, that is, for all  $i \in \{A, B\}$  and  $k \in \{1, 2\}$ ,  $u_i(\cdot, \theta^k) \equiv u(\cdot, \theta^k)$ . Assume also w.l.o.g. that  $u(0, R^0, \theta^1) = u(0, R^0, \theta^2) \equiv U^0$ . We will call to this value  $U^0$  the reservation utility level. For all  $i \in \{A, B\}$ , all  $\theta_i \in \{\theta^1, \theta^2\}$ , and all  $c \in \mathbb{R}_+^2$ , we say that contract  $c$  is *individually rational* for agent  $i$  when  $u(c, \theta_i) \geq U^0$ .

We will suppose that, for all  $\theta^k \in \{\theta^1, \theta^2\}$ , the utility functions  $u(\cdot, \theta^k)$  satisfy the following assumptions:

**Assumption 1.** For all  $\theta^k \in \{\theta^1, \theta^2\}$ ,  $u(\cdot, \theta^k)$  is twice continuous and differentiable in  $(x, R)$ , decreasing in  $x$ , increasing in  $R$  and strictly concave in  $(x, R)$ .

**Assumption 2** (*Single-crossing property*). For all  $(x, R) \in \mathbb{R}_+^2$  and all  $x' >$

---

<sup>2</sup>Notice that this model does not include joint production processes where the effort of both agents is necessary to produce a unique output.

$x$ , let  $R_1, R_2 \in \mathbb{R}_+$  be such that  $u(x, R, \theta^1) = u(x', R_1, \theta^1)$  and  $u(x, R, \theta^2) = u(x', R_2, \theta^2)$ . Then  $R_1 > R_2$ .

From these assumptions it follows that preferences on  $\mathbb{R}_+^2$  are strictly convex and that the indifference curves of the less and the more productive agent only cross in one point<sup>3</sup>.

Suppose that the principal could observe the type of the agents before signing any contract (i.e. he had complete information). Then, for all  $i \in \{A, B\}$  and all  $k \in \{1, 2\}$ , if agent  $i$  is of type  $k$ , the principal would offer him the following contract:

$$c_k^* = (x_k^*, R_k^*) \in \arg \max_{c \in \mathbb{R}_+^2} \{\pi_c : u(c, \theta^k) \geq U^0\} \quad (1)$$

The solutions to these programs for all  $k \in \{1, 2\}$  are what we call first-best contracts. We will assume that these contracts are such that  $\pi_{c_1^*} > 0$  and  $\pi_{c_2^*} > 0$ . It is easy to see that the first best contracts are defined by the following two equations:

$$u(c_k^*, \theta^k) = U^0 \quad (2)$$

$$\frac{\partial u(x_k^*, R_k^*, \theta^k)}{\partial R} = -\frac{\partial u(x_k^*, R_k^*, \theta^k)}{\partial x} \quad (3)$$

and that they verify the following properties:

$$\pi_{c_1^*} < \pi_{c_2^*} \quad (4)$$

$$u(c_2^*, \theta^1) < U^0 < u(c_1^*, \theta^2) \quad (5)$$

We have represented graphically the first-best contracts in Figure 2.1. Here, solid (dotted) curves represent indifference curves of an agent of type 1 (type 2). It is clear that any individually rational contract for an agent of type 1 is also individually rational for an agent of type 2. Straight lines with slope one represent contracts which give to the principal the same profits.

---

<sup>3</sup>Most papers studying this problem assume that preferences can be represented by an additively separable utility function,  $u(x, R, \theta_i) = U(R) - D(x, \theta_i)$ , where  $U$  is increasing and strictly concave,  $D$  is increasing and strictly convex in  $x$ , and for all  $x$ ,  $D(x, \theta^1) > D(x, \theta^2)$  and  $D_x(x, \theta^1) > D_x(x, \theta^2)$ . These functions are included in our domain.

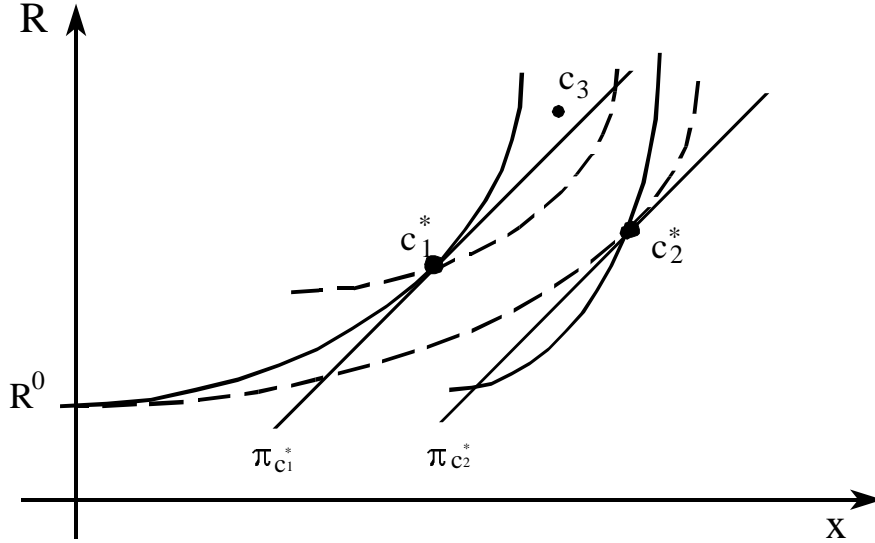


Figure 2.1:

In order to be sure that the agents will accept to sign the first-best contracts we assume that they strictly prefer to sign a contract which gives them the reservation utility level, rather than not signing any contract<sup>4</sup>. From now, we denote as  $NC$  to the situation in which an agent signs no contract. Then we can write our assumption as follows:

**Assumption 3.** For all  $i \in \{A, B\}$  and all  $c \in \mathbb{R}_+^2$  with  $u(c, \theta_i) \geq U^0$ , agent  $i$  strictly prefers  $c$  rather than  $NC$ . Moreover, for all  $c' \in \mathbb{R}_+^2$  with  $u(c', \theta_i) < U^0$ , agent  $i$  strictly prefers  $NC$  rather than  $c'$ .

Under Assumption 3, it is clear that we can not identify the situation  $NC$  with any contract in  $\mathbb{R}_+^2$ . To represent the preference relation over the whole set  $\mathbb{R}_+^2 \cup NC$  we define a utility function  $v(\cdot, \theta_i) : \mathbb{R}_+^2 \cup NC \rightarrow \mathbb{R}$  such that, (i) for all

---

<sup>4</sup>For that, we can assume that the agents do not like to be unemployed, and that then, they prefer to have a job that allows them to obtain the reservation utility level rather than to be out of work and to be paid the unemployment compensation (although they still prefer to be unemployed rather than to sign a contract which gives them an utility level smaller than  $U^0$ ).



$c \in \mathbb{R}_+^2$  with  $u(c, \theta_i) \geq U^0$ ,  $v(c, \theta_i) = u(c, \theta_i) + \varepsilon$  (where  $\varepsilon > 0$ ), (ii) for all  $c \in \mathbb{R}_+^2$  with  $u(c, \theta_i) < U^0$ ,  $v(c, \theta_i) = u(c, \theta_i)$ , and (iii)  $v(NC, \theta_i) = U^0$ .

Now, we introduce some additional notation. A *state of the world* is a profile  $(\theta_A, \theta_B) \in \{\theta^1, \theta^2\} \times \{\theta^1, \theta^2\}$ . For all  $k, l \in \{1, 2\}$  we denote  $\theta^{kl}$  to the state of the world  $(\theta_A, \theta_B) = (\theta^k, \theta^l)$ . Let  $\Theta = \{\theta^{11}, \theta^{12}, \theta^{21}, \theta^{22}\}$  be the *set of feasible states of the world*. For all  $\theta^{kl} \in \Theta$ , all  $c = (x, R) \in \mathbb{R}_+^2$  and all  $i \in \{A, B\}$ , let  $v_i(c, \theta^{kl})$  be the utility level obtained by agent  $i$  with contract  $c$  given that the state of the world is  $\theta^{kl}$  (i.e.  $v_A(c, \theta^{kl}) \equiv v(c, \theta^k)$  and  $v_B(c, \theta^{kl}) \equiv v(c, \theta^l)$ ). In the same way,  $v_i(NC, \theta^{kl})$  represents the utility level of agent  $i$  if he does not sign any contract and the state of the world is  $\theta^{kl}$  (i.e.  $v_i(NC, \theta^{kl}) \equiv U^0$ ).

Let  $C = \{c = (x, R) \in \mathbb{R}_+^2 : x - R \geq 0\}$  be the *set of feasible contracts*. The *set of feasible outcomes* that can be offered by the principal is  $C \cup NC$ . A *choice rule* is a mapping  $\varphi : \Theta \rightarrow (C \cup NC) \times (C \cup NC)$ , which associates each state of the world with a feasible outcome for each agent. Specifically, let  $\Phi$  be the choice rule that selects for each state of the world the first best contracts (i.e. for all  $\theta^{kl} \in \Theta$ ,  $\Phi(\theta^{kl}) = (c_k^*, c_l^*)$ ). We call this rule the *First Best Choice Rule (FBCR)*.

A *mechanism* is a tuple  $\Gamma = (S, g)$  where for all  $i \in \{A, B\}$ ,  $S_i$  is the strategy space for agent  $i$ ,  $S = S_A \times S_B$ , and  $g : S \rightarrow (C \cup NC) \times (C \cup NC)$  is an outcome function which associates to each profile of strategies some feasible outcomes. For all profile of strategies  $s \in S$  and all  $i \in \{A, B\}$ ,  $s_i$  and  $s_{-i}$  denote the strategies of agent  $i$  and the strategy of the other agent respectively. For all  $s \in S$  and all  $i \in \{A, B\}$ , we denote as  $g_i(s)$  the outcome selected by the mechanism for agent  $i$  when agents announce  $s$ . For all  $\theta^{kl} \in \Theta$ , all  $s \in S$  and all  $i \in \{A, B\}$ ,  $v_i(g(s), \theta^{kl})$  denote the utility level obtained by  $i$  with the outcome that the mechanism select for him given that the agents have announced  $s$  and the state of the world is  $\theta^{kl}$  (i.e.  $v_i(g(s), \theta^{kl}) \equiv v_i(g_i(s), \theta^{kl})$ ).

A *concept of equilibrium*,  $E$ , is a mapping from  $\Theta$  to  $S$ ,  $E : \Theta \rightarrow S$ . In particular, the Nash equilibrium is defined as follows:

**Definition 1.** Let  $\Gamma = (S, g)$  be a mechanism. We say that  $(s_A, s_B) \in S$  is a **Nash equilibrium** of mechanism  $\Gamma$  for the state of the world  $\theta^{kl}$  when, for all  $i \in \{A, B\}$  and all  $s'_i \in S_i$ ,  $v_i(g(s_A, s_B), \theta^{kl}) \geq v_i(g(s'_i, s_{-i}), \theta^{kl})$ .

For all concept of equilibrium  $E$  and all state of the world  $\theta^{kl} \in \Theta$ , let  $E(\Gamma, \theta^{kl})$  be the set of equilibrium strategies of mechanism  $\Gamma$  when the state of the world is  $\theta^{kl}$ . Specifically,  $N(\Gamma, \theta^{kl})$  denotes the set of Nash equilibrium strategies. We say that mechanism  $\Gamma$  *implements* the choice rule  $\varphi$  in  $E$ -equilibrium when, for all  $\theta^{kl} \in \Theta$ ,  $g(E(\Gamma, \theta^{kl})) = \varphi(\theta^{kl}) \neq \emptyset$ .

Since we are assuming that the principal does not observe the realization of  $(\theta_A, \theta_B)$ , his aim is then to design a mechanism that implements the FBCR.

From Equation 5 we have that, if agent  $i$  is more productive ( $\theta^2$ ), he strictly prefers the contract designed for a less productive type. For this reason, in any direct mechanism in which each agent only has to announce his type and where the outcome function gives each agent the first-best contract corresponding to his announcement, truth telling is not a dominant strategy for some states of the world. A well-known result in the literature (the Revelation Principle, see Gibbard [5]) tell us that then, the FBCR can not be implemented in dominant strategies. This lead us to focus on implementation in Nash equilibrium.

### 3. DESIRABLE PROPERTIES FOR THE IMPLEMENTING MECHANISMS

Maskin [11] showed that any choice rule which is Nash implementable must satisfy a monotonicity condition. Roughly speaking, this condition says that if the choice rule selects some outcome for some preferences profile, it must select the same outcome if the preferences of all agents change in such a way that this outcome has moved up in everyone rankings. Notice that the FBCR satisfies this condition in an obvious way: given any preferences profile and the outcome selected by the FBCR for it, there is no other preferences profile such that this outcome has moved up in the ranking of both agents. If there are three or more agents, this monotonicity condition together with an additional requirement called no veto power<sup>5</sup> are sufficient for implementation in Nash equilibrium. For the case in which there are only two agents, Moore and Repullo [13], and Dutta and Sen [4] have established a necessary and sufficient condition for Nash implementation.

In the proofs of these general results the authors propose some mechanisms which work for any choice rule. However, these mechanisms are not very appealing and they have been the object of criticism in the literature. One of the main objections against them is that they use integer games (i.e. games in which the agent who announces the highest integer obtains his most preferred outcome)<sup>6</sup>.

This indicates that such mechanisms should be taken just as a tool in the proof of general results. When we face up to a specific problem, we have to look for

---

<sup>5</sup>No veto power means that, if there exists an outcome which is the most preferred for all but one agent, it must be selected by the choice rule.

<sup>6</sup>See Jackson [7] and Moore [12] for a criticism of integer games.

specific mechanisms whose behavior is better within this particular framework. That is what we try to do in this paper: we study the possibility of implementing the FBCR in Nash equilibrium by means of mechanisms satisfying desirable properties. In particular, we will try to avoid mechanisms which use integer games. Nevertheless, this is not our only aim. Now we detail the properties that we would like to be satisfied by any mechanism implementing the FBCR in Nash equilibrium.

### 3.1. Strategy Spaces as Small as Possible

We try to avoid mechanisms with large strategy spaces because they usually make it difficult to understand their rules to the agents. Moreover, the larger are the strategy spaces, the more difficult is that agents coordinate among them to play the equilibrium strategies. The next two propositions tell us the minimum size of the strategy spaces of any mechanism implementing the FBCR in Nash equilibrium.

**Proposition 1.** *Let  $\Gamma = (S, g)$  be a mechanism implementing the FBCR in Nash equilibrium. Then, at least for one agent  $i \in \{A, B\}$ , the strategy space  $S_i$  is composed by more than three elements.*

**Proof.** Let  $\Gamma = (S, g)$  be a mechanism that implements the FBCR in Nash equilibrium and suppose by contradiction that, for all  $i \in \{A, B\}$ ,  $S_i$  has as much as three elements. For each  $\theta^{kl} \in \Theta$ , let  $(s_A^{kl}, s_B^{kl}) \in N(\Gamma, \theta^{kl})$ . Obviously,  $g(s_A^{kl}, s_B^{kl}) = (c_k^*, c_l^*)$ . Notice that, for all  $i \in \{A, B\}$ ,  $s_i^{11} \neq s_i^{22}$  (otherwise  $(s_A^{22}, s_B^{22}) \notin N(\Gamma, \theta^{22})$ ). Moreover,  $s_A^{12} \neq s_A^{11}$  (otherwise  $(s_A^{12}, s_B^{12}) \notin N(\Gamma, \theta^{12})$ ), and  $s_B^{12} \neq s_B^{22}$  (otherwise  $(s_A^{22}, s_B^{22}) \notin N(\Gamma, \theta^{22})$ ). Suppose that  $s_A^{12} \neq s_A^{22}$  and  $s_B^{12} \neq s_B^{11}$ . Then, since for all  $i \in \{A, B\}$ ,  $S_i$  has as much three elements, we have that, for all  $i \in \{A, B\}$ ,  $S_i = \{s_i^{11}, s_i^{22}, s_i^{12}\}$ . Therefore,  $(s_A^{21}, s_B^{21}) = (s_A^{11}, s_B^{22})$  (otherwise we have that, for some  $\theta^{kl} \in \Theta$ ,  $(s_A^{kl}, s_B^{kl}) \notin N(\Gamma, \theta^{kl})$ ). For all  $i \in \{A, B\}$ , all  $\theta_i \in \{\theta^1, \theta^2\}$  and all  $s_{-i} \in S_{-i}$ , let  $s_i(\theta_i/s_{-i}) \equiv \arg \max\{v(g_i(s_i, s_{-i}), \theta_i) : s_i \in S_i\}$ . Step 1. Since  $g_B(s_A^{11}, s_B^{11}) = g_B(s_A^{11}, s_B^{22}) = c_1^*$  and  $(s_A^{11}, s_B^{11}) \notin N(\Gamma, \theta^{12})$ , then  $\frac{s_B(\theta^2/s_A^{11})}{s_A^{11}} = s_B^{12}$ . Moreover,  $v(g_B(s_A^{11}, s_B(\theta^2/s_A^{11})), \theta^2) > v(c_1^*, \theta^2)$ , and then we have that  $g_B(s_A^{11}, s_B^{12}) \neq c_2^*$ . Step 2. Since  $(s_A^{12}, s_B^{12}) \notin N(\Gamma, \theta^{22})$ , it is clear that  $s_A(\theta^2/s_B^{12}) \neq s_A^{12}$ . Moreover, since  $s_B(\theta^2/s_A^{11}) = s_B^{12}$ , then  $s_A(\theta^2/s_B^{12}) \neq s_A^{11}$  (otherwise  $(s_A^{11}, s_B^{12}) \in N(\Gamma, \theta^{22})$ , which is a contradiction, since  $g_B(s_A^{11}, s_B^{12}) \neq c_2^*$ ). Therefore  $s_A(\theta^2/s_B^{12}) = s_A^{22}$ .

Moreover, since  $(s_A^{12}, s_B^{12}) \notin N(\Gamma, \theta^{22})$ , we have that  $v(g_A(s_A(\theta^2/s_B^{12}), s_B^{12}), \theta^2) > v(c_1^*, \theta^2)$ , and then  $g_A(s_A^{22}, s_B^{12}) \neq c_2^*$ .

Step 3. Since  $g_A(s_A^{22}, s_B^{22}) = g_A(s_A^{11}, s_B^{22}) = c_2^*$  and  $(s_A^{22}, s_B^{22}) \notin N(\Gamma, \theta^{12})$ , then  $s_A(\theta^1/s_B^{22}) = s_A^{12}$ . Moreover, since  $(s_A^{22}, s_B^{22}) \in N(\Gamma, \theta^{22})$  and  $v(c_1^*, \theta^2) > v(c_2^*, \theta^2)$ , we have that  $g_A(s_A^{12}, s_B^{22}) \neq c_1^*$ .

Step 4. Since  $(s_A^{12}, s_B^{12}) \notin N(\Gamma, \theta^{11})$ , it is clear that  $s_B(\theta^1/s_A^{12}) \neq s_B^{12}$ . Moreover, since  $s_A(\theta^1/s_B^{22}) = s_A^{12}$ , then  $s_B(\theta^1/s_A^{12}) \neq s_B^{22}$  (otherwise  $(s_A^{12}, s_B^{22}) \in N(\Gamma, \theta^{11})$ , which is a contradiction, since  $g_A(s_A^{12}, s_B^{22}) \neq c_1^*$ ). Therefore  $s_B(\theta^1/s_A^{12}) = s_B^{11}$ . Moreover, since  $(s_A^{12}, s_B^{12}) \in N(\Gamma, \theta^{12})$  and  $v(c_1^*, \theta^2) > v(c_2^*, \theta^2)$ , we have that  $g_B(s_A^{12}, s_B^{11}) \neq c_1^*$ .

Step 5. Since  $(s_A^{11}, s_B^{11}) \notin N(\Gamma, \theta^{21})$ , it is clear that  $s_A(\theta^2/s_B^{11}) \neq s_A^{11}$ . Moreover, since  $s_B(\theta^1/s_A^{12}) = s_B^{11}$ , then  $s_A(\theta^2/s_B^{11}) \neq s_A^{12}$  (otherwise  $(s_A^{12}, s_B^{11}) \in N(\Gamma, \theta^{21})$ , which is a contradiction, since  $g_B(s_A^{12}, s_B^{11}) \neq c_1^*$ ). Therefore  $s_A(\theta^2/s_B^{11}) = s_A^{22}$ . Moreover, since  $(s_A^{11}, s_B^{11}) \notin N(\Gamma, \theta^{21})$ , we have that  $v(g_A(s_A(\theta^2/s_B^{11}), s_B^{11}), \theta^2) > v(c_1^*, \theta^2)$ , and then  $g_A(s_A^{22}, s_B^{11}) \neq c_2^*$ .

Step 6. Since  $(s_A^{22}, s_B^{22}) \notin N(\Gamma, \theta^{21})$ , it is clear that  $s_B(\theta^1/s_A^{22}) \neq s_B^{22}$ . Moreover, since  $s_A(\theta^2/s_B^{11}) = s_A^{22}$ , then  $s_B(\theta^1/s_A^{22}) \neq s_B^{11}$  (otherwise  $(s_A^{22}, s_B^{11}) \in N(\Gamma, \theta^{21})$ , which is a contradiction, since  $g_A(s_A^{22}, s_B^{11}) \neq c_2^*$ ). Finally, since  $s_A(\theta^2/s_B^{11}) = s_A^{22}$ , then  $s_B(\theta^1/s_B^{22}) \neq s_B^{12}$  (otherwise  $(s_A^{22}, s_B^{12}) \in N(\Gamma, \theta^{21})$ , which is a contradiction, since  $g_A(s_A^{22}, s_B^{12}) \neq c_2^*$ ).

From Step 6 we have that  $s_B(\theta^1/s_A^{22}) = \emptyset$ , which is a contradiction. The cases where  $s_A^{12} = s_A^{22}$  or  $s_B^{12} = s_B^{11}$  are similar. ■

**Proposition 2.** *Let  $\Gamma = (S, g)$  be a mechanism implementing the FBCR in Nash equilibrium such that, at least for one agent  $i \in \{A, B\}$ , the strategy space  $S_i$  is composed by only four elements. Then, the strategy space of the other agent  $j \in \{A, B\}$ ,  $j \neq i$ , is composed by more than two elements.*

The proof of this proposition is similar to the one of Proposition 1 and we omit it. From Propositions 1 and 2 we have that, in any mechanism implementing the FBCR in Nash equilibrium, at least the strategy space of one agent has four elements and the strategy space of the other agent has three elements<sup>7</sup>.

---

<sup>7</sup>The issue of the dimensionality of the strategy space has been studied by Reichelstein and Reiter [14] among others. They only consider the case in which strategy spaces are Euclidean and thus, a natural measure of size is dimensionality. In our case, given the nature of the mechanisms that we will propose, the natural measure of size is the number of elements of the strategy space.

### 3.2. Enforceability

An important issue to be taken into consideration when we design a mechanism implementing the FBCR is whether the agents who play the mechanism are really obliged to accept the contracts that the outcome function chooses. This fact will be of fundamental importance in order to know the way in which agents work out their expected utility levels associated to each profile of strategies (and therefore, to know which are the equilibria of the game associated with the mechanism). If we were able to force agents to sign the contracts that the mechanism selects for each profile of strategies that they play (whatever it is), then agents' expected utility with these strategies will only depend on such contracts. On the contrary, if they may reject the outcome of the mechanism, no agent will accept any contract which is not individually rational for him. In that case, the utility level that an agent gets will not be the one for the non-individually rational contract that the mechanism chooses, but the reservation utility level.

Ideally, we would allow agents to reject contracts and we would design a mechanism that implements the FBCR taking into account that if we give a non-individually rational contract to some agent, he will obtain the reservation utility level. Such mechanism could include contracts of three categories: (1) contracts that are individually rational for none type of agent, and that then would never be accepted (that is, contracts which are non-individually rational for an agent of type 2, since such contracts are non-individually rational for an agent of type 1 as well), (2) contracts that are individually rational for both types of agents, and that then, would always be accepted (that is, contracts that are individually rational for an agent of type 1, since they would be individually rational for an agent of type 2 as well) and, (3) contracts that are individually rational only for agents of type 2, and that then, sometimes would be accepted and sometimes would be rejected (for example contract  $c_3$  in Figure 5.1, which is individually rational for an agent of type 2, but it is not individually rational for an agent of type 1).

Unfortunately, as the next proposition shows, if agents may walk away from the mechanism we can not implement in Nash equilibrium the FBCR.

**Proposition 3.** *If agents may refuse to sign the contracts chosen by the mechanism then the FBCR can not be implemented in Nash equilibrium.*

**Proof.** Let  $\Gamma = (S, g)$  be a mechanism that implements the FBCR in Nash equilibrium and suppose by contradiction that agents may refuse to sign the

contracts that the mechanism chooses. Notice that, for all  $i \in \{A, B\}$ , all  $\theta_i \in \{\theta^1, \theta^2\}$ , and all  $s \in S$  with  $v(g_i(s), \theta_i) < U^0$ , agent  $i$  will reject contract  $g_i(s)$  and then his expected utility from that profile of strategies will be  $U^0$ . Let  $(s_A^{22}, s_B^{22}) \in N(\Gamma, \theta^{22})$ . We have that for all  $i \in \{A, B\}$  and all  $s_i \in S_i$ ,  $v(g_i(s_i, s_{-i}^{22}), \theta^2) \leq v(g_i(s_A^{22}, s_B^{22}), \theta^2) = v(c_2^*, \theta^2)$ , and therefore  $v(g_i(s_i, s_{-i}^{22}), \theta^1) < U^0$ . This implies that for all  $i \in \{A, B\}$  with  $\theta_i = \theta^1$  and all  $s_i \in S_i$ , the agent  $i$ 's expected utility from strategies  $(s_i, s_{-i}^{22})$  is  $U^0$ . Therefore  $(s_A^{22}, s_B^{22})$  is a Nash equilibrium of mechanism  $\Gamma$  when the state of the world is  $\theta^{11}$ , which is a contradiction. ■

This negative result is due to the fact that when agents are of type 1 there are Nash equilibria in which they do not sign any contract. It seems that, in order to implement the FBCR in Nash equilibrium, we have to force the agents to sign the resulting contracts of the mechanism. For that we have to think about the mechanism design as a three-stage problem. In the first stage the principal proposes a mechanism. In the second stage the agents decide whether they want to commit themselves to play the mechanism. Finally, in the third stage those agents who accepted the mechanism play the game induced by it. We will assume that agents anticipate that the outcome of the mechanism will be the equilibrium one, and therefore they will commit themselves to play any mechanism whose equilibrium contracts give them a utility level equal or greater than reservation utility level (and in particular, any mechanism implementing the FBCR).

However, this approach creates a new problem. It is clear that any mechanism implementing the FBCR will always select non-individually rational contracts for some type of agent (for example, the mechanism have to choose contract  $c_2^*$  for some strategies, and such contract is not individually rational for an agent of type 1). The question is then: even though an agent has committed himself to play the mechanism, can he be forced to sign the contract that the mechanism chooses in case it was non-individually rational for him?

It seems reasonable that, if an agent is able to prove in court that the contract that has been chosen by the mechanism is not individually rational for him, then the principal can not make him sign that contract (just as we can not force anyone to be a slave, although he had committed himself to be one). When will an agent be able to prove that a contract is not individually rational for him?

Suppose that agents can always prove when a contract is non-individually rational for him. This would be equivalent to say that an agent type 1 can prove what his type is (he just has to prove that a contract like  $c_3$  in Figure 5.1 is non-individually rational for him). In that case the problem of implementing the

FBCR would be obvious. All we have to do is to ask each agent to prove that he is type 1. If he prove that, then we offer him contract  $c_1^*$ , otherwise we offer him contract  $c_2^*$  (notice that, under Assumption 3, for an agent type 1 would be a dominant strategy to prove what his type is). Therefore, this case has no interest.

However, the logical thing is that agents can not prove what their types are. In this case, the only contracts that an agent could refuse to sign (even though they had been chosen by the mechanism that the agent promised to accept) would be the ones that are not individually rational for both types of agents. An impartial jury would be able to understand that such contracts are not individually rational for any agent (although they do not know their types). This, together with Proposition 3, leads us to consider that it would be desirable that our mechanisms never select contracts like these. In this sense, we have the following definition:

**Definition 2.** *We say that a mechanism  $\Gamma = (S, g)$  is **enforceable** if for all  $s \in S$  and all  $i \in \{A, B\}$ , there is some  $\theta_i \in \{\theta^1, \theta^2\}$  with  $v(g_i(s), \theta_i) \geq U^0$*

As Figure 2.1 makes clear, under our assumptions the later property is equivalent to saying that for all  $i \in \{A, B\}$  and all  $s \in S$ ,  $v(g_i(s), \theta^i) \geq U^0$ . As we have argued, if a mechanism implementing the FBCR does not satisfy this property, then it achieves its objective with threats which are not credible.

### 3.3. Renegotiation Proofness Among Agents

When an agent  $i \in \{A, B\}$  signs contract  $(x_i, R_i) \in C$  with the principal, they are establishing a double commitment. On the one hand, agent  $i$  is committing himself to produce an amount of output  $x_i$  for the principal. On the other hand, the principal is committing himself to pay an amount  $R_i$  to agent  $i$ . Suppose now that once the principal has signed a contract  $(x_i, R_i) \in C$  with each agent  $i \in \{A, B\}$ , agents decide to renegotiate these contracts among them. For example, the agents could reach an agreement in such a way that agent  $A$  gives to agent  $B$  part of the payoff that he will receive from the principal and, in exchange for that, agent  $B$  has to produce part of the output that agent  $A$  had committed to produce for the principal. Obviously, any of such renegotiations have to be such that, at the end: (i) the sum of the quantities produced by both agents equals the sum of the quantities that each agent separately had committed himself to produce, and (ii) the sum of the payoffs that each agent really receives equals the sum of the payoffs that the principal had committed himself to give to each agent separately.

It is unlikely that the principal can prevent this kind of renegotiation among the agents: first, the principal may not know whether the output that each agent is giving to him has been really produced by this agent, and secondly, the principal can not control what agents do with their income. Agents will have incentives to carry out such renegotiations when, by doing that, they end better off (compared with the contracts that they really signed). Notice that the way in which the agents share out some given contracts may depends on their types (for example, the same renegotiation of some given profile of contracts could be beneficial for an agent when he is of type 1, and to be damaging for him when he is of type 2).

Renegotiation changes the strategic aspect of the problem of designing a mechanism. Now, agent  $i$ 's utility level associated to some profile of strategies is not determined by the contract that the mechanism selects for him, but it is determined by the final redistribution of the profile of contracts selected by the mechanism. Then, profiles of contracts that were not desirable for some agent may become more attractive for him if renegotiation is considered. A mechanism implementing the FBCR should take this into account.

Now we introduce some notation. Let  $\bar{c} \in (C \cup NC) \times (C \cup NC)$  be a profile of feasible outcomes. A *renegotiation rule* will be a mapping  $F : (C \cup NC) \times (C \cup NC) \times \Theta \rightarrow (\mathbb{R}_+^2 \cup NC) \times (\mathbb{R}_+^2 \cup NC)$ , such that:

(i) for all  $\bar{c} = (c_A, c_B) \in C \times C$  and all  $\theta^{kl} \in \Theta$ ,  $F(\bar{c}, \theta^{kl}) = (F_A(\bar{c}, \theta^{kl}), F_B(\bar{c}, \theta^{kl})) \in \mathbb{R}_+^2 \times \mathbb{R}_+^2$  satisfies that  $F_A(\bar{c}, \theta^{kl}) + F_B(\bar{c}, \theta^{kl}) = c_A + c_B$  (where, for all  $i \in \{A, B\}$ ,  $F_i(\bar{c}, \theta^{kl})$  denotes the output and the payoff corresponding to agent  $i$  after the renegotiation of the contracts  $\bar{c}$  given that the state of the world is  $\theta^{kl}$ ), and

(ii) for all  $c \in C$  and all  $\theta^{kl} \in \Theta$ ,  $F((c, NC), \theta^{kl}) = (c, NC)$ ,  $F((NC, c), \theta^{kl}) = (NC, c)$  and  $F((NC, NC), \theta^{kl}) = (NC, NC)$  (that is, if there is one agent who does not sign any contract then no renegotiation is possible).

For all state of the world  $\theta^{kl} \in \Theta$ , all profile of feasible outcomes  $\bar{c} \in (C \cup NC) \times (C \cup NC)$ , all agent  $i \in \{A, B\}$ , and all renegotiation rule  $F$ , let  $v_i(F(\bar{c}), \theta^{kl})$  be the utility obtained by agent  $i$  after renegotiate according to  $F$  the outcomes  $\bar{c}$  when the state of the world is  $\theta^{kl}$  (i.e.  $v_i(F(\bar{c}), \theta^{kl}) \equiv v_i(F_i(\bar{c}, \theta^{kl}), \theta^{kl})$ ).

We will say that a renegotiation rule  $F$  is *admissible* when for all  $i \in \{A, B\}$ , all  $\bar{c} \in C \times C$  and all  $\theta^{kl} \in \Theta$ ,  $v_i(F(\bar{c}), \theta^{kl}) \geq v_i(\bar{c}, \theta^{kl})$ . That is, we do not admit renegotiation rules where some agent is worse after the renegotiation than with his initial contract (this can be interpreted as a requirement of individual rationality on the renegotiation). Let  $\mathcal{F}$  be the *class of admissible renegotiation*



rules. Notice that for all  $F \in \mathcal{F}$ , all  $\bar{c} = (c_A, c_B) \in C \times C$  and all  $\theta^{kl} \in \Theta$ ,  $\pi_{c_A} + \pi_{c_B} = \pi_{F_A(\bar{c}, \theta^{kl})} + \pi_{F_B(\bar{c}, \theta^{kl})}$ .

Now, we have the following definition:

**Definition 3.** For all mechanism  $\Gamma = (S, g)$ , all admissible renegotiation rule  $F \in \mathcal{F}$  and all state of the world  $\theta^{kl} \in \Theta$ , we say that  $s \in S$  is a **F-Nash equilibrium** of  $\Gamma$  for state of the world  $\theta^{kl}$  when, for all  $i \in \{A, B\}$  and all  $s'_i \in S_i$ ,  $v_i(F(g(s)), \theta^{kl}) \geq v_i(F(g(s'_i, s_{-i})), \theta^{kl})$ . Let  $F-N(\Gamma, \theta^{kl})$  denote the set of  $F$ -Nash equilibrium of mechanism  $\Gamma$  when the state of the world is  $\theta^{kl}$ .

In words, a profile of strategies is a  $F$ -Nash equilibrium when, taking into account that the renegotiation rule used by the agents is  $F$ , no agent can improve by means of an unilateral deviation.

Let  $F^I$  denotes the situation where the agents do not renegotiate their contracts. It is clear that  $F^I \in \mathcal{F}$ , and that the notion of  $F^I$ -Nash equilibrium matches the usual definition of Nash equilibrium. For simplicity, we will still write Nash equilibrium when we are talking about a  $F^I$ -Nash equilibrium (and, for all mechanism  $\Gamma = (S, g)$  and all  $\theta^{kl} \in \Theta$ , we will still write  $N(\Gamma, \theta^{kl})$  instead of  $F^I-N(\Gamma, \theta^{kl})$ ).

Suppose that the principal knows that the agents will renegotiate their contracts in accordance with some given renegotiation rule  $F \in \mathcal{F}$ . In that case he have to design a mechanism implementing the FBCR not in Nash equilibrium, but in  $F$ -Nash equilibrium.

Obviously, given any  $F \in \mathcal{F}$ , a mechanism implementing the FBCR in  $F$ -Nash equilibrium may not work for any other different renegotiation rule (i.e. given any  $\tilde{F} \in \mathcal{F}$  with  $\tilde{F} \neq F$ , the former mechanism may not implement the FBCR in  $\tilde{F}$ -Nash equilibrium). Then, what happens if the principal does not know the specific renegotiation rule that will be used by the agents, or if he does not even know whether agents will renegotiate their contracts at the end?

The best thing would be to design a mechanism implementing the FBCR in  $F$ -Nash equilibrium whatever the renegotiation rule  $F \in \mathcal{F}$  is. If we are able to design a mechanism like that, then it does not matter that the principal does not know the renegotiation rule used by the agents. In this sense we have the following definition:

**Definition 4.** We say that a mechanism  $\Gamma = (S, g)$  is **renegotiation-proof** when, for all admissible renegotiation rule  $F \in \mathcal{F}$ ,  $\Gamma$  implements the FBCR in

$F$ -Nash equilibrium<sup>8</sup>.

### 3.4. Unique Implementation

Let  $F \in \mathcal{F}$  be the renegotiation rule used by the agents and let  $\Gamma = (S, g)$  be a mechanism that implements the FBCR in  $F$ -Nash equilibrium. Suppose that our implementing mechanism is such that for some state of the world there are at least two profiles of strategies which are a  $F$ -Nash equilibrium. It is clear that both profiles of strategies yield the same outcome: the first-best contracts. However, if one of the agents decides to play the strategy corresponding to the first profile of strategies and the other agent plays the strategy corresponding to the second one, the resulting outcome may not be the first-best contracts. The existence of more than one equilibrium for some state of the world creates a problem of coordination among agents. In order to implement the FBCR we have to be certain that agents coordinate their decisions to play the same equilibrium.

To avoid this problem we would like that our mechanisms had a unique equilibrium for each state of the world. For that, we have the following definition:

**Definition 5.** Let  $F \in \mathcal{F}$  be an admissible renegotiation rule and let  $\Gamma = (S, g)$  be a mechanism that implements the FBCR in  $F$ -Nash equilibrium. We say that  $\Gamma$  satisfies the **F-unique implementation property** when for all  $\theta^{kl} \in \Theta$ ,  $F - N(\Gamma, \theta^{kl})$  has only one element.

To continue with the notation of the previous subsection, we will write unique implementation property to denote the  $F^I$ -unique implementation property.

### 3.5. Unique Best-Reply

Let  $F \in \mathcal{F}$  be the renegotiation rule used by the agents and let  $\Gamma = (S, g)$  be a mechanism that implements the FBCR in  $F$ -Nash equilibrium. If a profile of strategies is a  $F$ -Nash equilibrium then no agent can improve by means of any unilateral deviation. For this reason we usually admit that, if any agent (say agent  $i$ ) thinks that the rest of agents will play the equilibrium strategies, he will play his equilibrium strategy too. However, what happen if there is another strategy for agent  $i$ , different from the equilibrium one, which gives him the same utility?

---

<sup>8</sup>Notice that, since  $F^I \in \mathcal{F}$ , any renegotiation-proof mechanism also implements the FBCR in Nash equilibrium.

In that case agent  $i$  will be indifferent between both strategies. This made us wonder whether agent  $i$  will play his equilibrium strategy.

To clear up this doubt we would like that, in case that the rest of agents play their equilibrium strategies, any strategy of agent  $i$  different from the equilibrium one gets him a level of utility strictly smaller. For that we make use of the following two definitions:

**Definition 6.** Given any renegotiation rule  $F \in \mathcal{F}$ , any mechanism  $\Gamma = (S, g)$ , and any state of the world  $\theta^{kl} \in \Theta$ , we say that  $s \in S$  is a **F-strict-Nash equilibrium** of mechanism  $\Gamma$  when the state of the world is  $\theta^{kl}$  when, for all  $i \in \{A, B\}$  and all  $s'_i \in S_i \setminus \{s_i\}$ ,  $v_i(F(g_i(s)), \theta^{kl}) > v_i(F(g_i(s'_i), s_{-i}), \theta^{kl})$ . Let  $F - SN(\Gamma, \theta^{kl})$  denote the set of these strict-Nash equilibrium.

**Definition 7.** Given any renegotiation rule  $F \in \mathcal{F}$ , let  $\Gamma = (S, g)$  be a mechanism implementing the FBCR in  $F$ -Nash equilibrium. We will say that  $\Gamma$  satisfies the **F-unique best-reply property** when, for all state of the world, any  $F$ -Nash equilibrium is also a  $F$ -strict Nash equilibrium (that is, for all  $\theta^{kl} \in \Theta$ ,  $F - SN(\Gamma, \theta^{kl}) = F - N(\Gamma, \theta^{kl})$ ).

Let  $F \in \mathcal{F}$  be the renegotiation rule used by the agents, and let  $\Gamma = (S, g)$  be a mechanism that implements the FBCR in  $F$ -Nash equilibrium. If the implementing mechanism satisfies the  $F$ -unique best-reply property, then the agents have strong incentives to stick to equilibrium strategies. Notice that if a mechanism implements the FBCR in  $F$ -Nash equilibrium and satisfies the  $F$ -unique best-reply property, then it doubly implements the FBCR in  $F$ -Nash and  $F$ -strict-Nash equilibrium. Although the inverse is not necessarily true, it will be if the mechanism also satisfies the  $F$ -unique implementation property.

Again, to continue with previous notation, we will write unique best-reply property to denote the  $F^I$ -unique best-reply property.

### 3.6. No Mixed Strategies

Let  $\Gamma = (S, g)$  be a mechanism. A *mixed strategy* for agent  $i$  is a probability distribution on  $S_i$ . Formally, a mixed strategy for agent  $i$  is a mapping  $\sigma_i : S_i \rightarrow [0, 1]$  such that  $\sum_{s_i \in S_i} \sigma_i(s_i) = 1$  (where  $\sigma_i(s_i)$  denotes the probability that agent  $i$  assigns to announce  $s_i$ ). For all  $i \in \{A, B\}$ , let  $\Sigma_i$  be the set of admissible probability distributions on  $S_i$ , and let  $\Sigma = \Sigma_A \times \Sigma_B$  be the set of admissible profiles of mixed strategies. For all  $\sigma_i \in \Sigma_i$ , let  $S_i(\sigma_i)$  be the set of elements of

$S_i$  that are announced with strictly positive probability according to strategy  $\sigma_i$ , that is  $S_i(\sigma_i) = \{s_i \in S_i : \sigma_i(s_i) > 0\}$ . We say that  $\sigma_i \in \Sigma_i$  is a *pure strategy* for agent  $i$  when, for some  $s_i \in S_i$ ,  $S_i(\sigma_i) = s_i$  (i.e. agent  $i$  announce with probability one some element of  $S_i$ ).

Up to now we have only considered the specific case in which agents play pure strategies. However, sometimes agents may find profitable to play another type of mixed strategies. Therefore, if we want to be sure that we are implementing the FBCR, our mechanisms should work not only for pure strategies, but also for mixed strategies.

Considering mixed strategies force us to generalize some of our previous definitions on implementation. First notice that when agents play mixed strategies we may not know exactly which will be the outcome selected by the mechanism. All that we will have is a probability distribution among all possible outcomes that can be selected by the mechanism. Formally, for all mechanism  $\Gamma = (S, g)$ , all  $(\sigma_A, \sigma_B) \in \Sigma$  and all  $(s_A, s_B) \in S$ , mechanism  $\Gamma$  will select the outcome  $g(s_A, s_B)$  with probability  $\sigma_A(s_A) \times \sigma_B(s_B)$ . Moreover, if the agents are using the renegotiation rule  $F \in \mathcal{F}$  and the state of the world is  $\theta^{kl} \in \Theta$ , then they will obtain the outcome  $F(g(s_A, s_B), \theta^{kl})$  with probability  $\sigma_A(s_A) \times \sigma_B(s_B)$ . Therefore, for knowing how good is a profile of mixed strategies for an agent, he has to calculate his expected utility. To be consistent with previous notation, for all  $F \in \mathcal{F}$ , all  $i \in \{A, B\}$ , all  $\theta^{kl} \in \Theta$ , and all  $\sigma \in \Sigma$ , let  $v_i^e(F(g(\sigma)), \theta^{kl})$  denote the expected utility<sup>9</sup> of agent  $i$  when agents play the profile of mixed strategies  $\sigma$ , the renegotiation rule is  $F$  and the state of the world is  $\theta^{kl}$ .

In this way, now a *concept of equilibrium*  $E$  will be a mapping from the set of feasible states of the world to the set of admissible profiles of mixed strategies,  $E : \Theta \rightarrow \Sigma$ . In particular, we have the following definition of  $F$ -Nash equilibrium when mixed strategies are considered:

**Definition 8.** For all mechanism  $\Gamma = (S, g)$ , all admissible renegotiation rule  $F \in \mathcal{F}$  and all state of the world  $\theta^{kl} \in \Theta$ , we say that  $\sigma \in \Sigma$  is a **F-Nash equilibrium** of  $\Gamma$  for state of the world  $\theta^{kl}$  when, for all  $i \in \{A, B\}$  and all  $\sigma'_i \in \Sigma_i$ ,  $v_i^e(F(g(\sigma)), \theta^{kl}) \geq v_i^e(F(g(\sigma'_i, \sigma_{-i})), \theta^{kl})$ . Let  $F - N(\Gamma, \theta^{kl})$  still denote the set of  $F$ -Nash equilibrium of mechanism  $\Gamma$  when the state of the world is  $\theta^{kl}$  (where  $F - N(\Gamma, \theta^{kl})$  may include now non-pure profiles of strategies).

In the same way that when we only considered pure strategies, we will say that a mechanism  $\Gamma = (S, g)$  implements the FBCR in  $F$ -Nash equilibrium when, for

---

<sup>9</sup>We will suppose that the usual von Neumann-Morgenstern assumptions hold.

all state of the world, the set of  $F$ -Nash equilibrium outcomes coincides with the first-best contracts (i.e. for all  $\theta^{kl} \in \Theta$ ,  $g(F - N(\Gamma, \theta^{kl})) = \Phi(\theta^{kl})$ ). Now that  $F - N(\Gamma, \theta^{kl})$  may include non-pure profiles of strategies, the natural way of defining the equilibrium outcomes is like those which have a strictly positive probability to occur given that the agents are playing a  $F$ -Nash equilibrium profile of strategies. That is,  $g(F - N(\Gamma, \theta^{kl})) = \{g(s_A, s_B) : \exists(\sigma_A, \sigma_B) \in F - N(\Gamma, \theta^{kl}) \text{ with } s_A \in S_A(\sigma_A) \text{ and } s_B \in S_B(\sigma_B)\}$ . Notice that then, since the FBCR is single valued, for all pair  $(s_A, s_B) \in S$  which is announced with strictly positive probability in some  $F$ -Nash equilibrium, the mechanism will select the same outcome: the first-best contracts. That is, for all  $\theta^{kl} \in \Theta$ , all  $(\sigma_A, \sigma_B) \in F - N(\Gamma, \theta^{kl})$ , all  $s_A \in S_A(\sigma_A)$  and all  $s_B \in S_B(\sigma_B)$ , we have that  $g(s_A, s_B) = (c_k^*, c_l^*)$ .

The existence of non-pure equilibrium strategies in our implementing mechanisms can create some difficulties. First, non-pure equilibrium strategies will be usually difficult to find. If we want that the agents find easily the solution to the game that we are proposing to them, we should try to avoid that this solution involves playing a non-pure strategy. Of course, one could argue that our implementing mechanisms could have pure and non-pure equilibrium strategies, and that in that case the agents would be able to find easily at least one of the solutions of the game. For that reason, we consider that a minimum requirement that any mechanism should satisfy is that, for all state of the world, there exists at least one equilibrium consisting of pure strategies.

**Definition 9.** Let  $F \in \mathcal{F}$  be an admissible renegotiation rule and let  $\Gamma = (S, g)$  be a mechanism implementing the FBCR in  $F$ -Nash equilibrium. We say that  $\Gamma$  satisfies the **F-pure equilibrium property** when, for all  $\theta^{kl} \in \Theta$ , there exists some  $(\sigma_A, \sigma_B) \in F - N(\Gamma, \theta^{kl})$  such that, for all  $i \in \{A, B\}$ ,  $\sigma_i$  is a pure strategy.

However, even in the case that our mechanisms satisfy this property, the existence of non-pure equilibrium strategies is still problematic. On the one hand, notice that the existence of pure and non-pure equilibrium strategies would contradict the spirit of the  $F$ -unique implementation property. On the other hand, the existence of non-pure equilibrium strategies also conflicts with the spirit of the  $F$ -unique best-reply property. To see that consider the following example. Let  $F \in \mathcal{F}$  be the renegotiation rule used by the agents, and let  $\Gamma = (S, g)$  be a mechanism implementing the FBCR in  $F$ -Nash equilibrium. Given any  $\theta^{kl} \in \Theta$ , let  $(\sigma_A, \sigma_B) \in F - N(\Gamma, \theta^{kl})$  be such that, for some  $i \in \{A, B\}$ ,  $\sigma_i$  is not a pure strategy. Then there exists at least two elements  $s_i, s'_i \in S_i(\sigma_i)$ . Let now  $\sigma'_i \in \Sigma_i$  be such that  $S_i(\sigma'_i) = s'_i$ . Notice that when the agents play

the profile of strategies  $(\sigma_i, \sigma_{-i})$  or  $(\sigma'_i, \sigma_{-i})$ , the mechanism will select with probability equal to one the same outcome,  $(c_k^*, c_i^*)$ . Therefore we have that  $v_i^e(F(g(\sigma_i, \sigma_{-i})), \theta^{kl}) = v_i^e(F(g(\sigma'_i, \sigma_{-i})), \theta^{kl})$ , and so  $\sigma_i$  is not the unique best-reply to  $\sigma_{-i}$ .

Therefore, given any  $F \in \mathcal{F}$ , if our mechanisms satisfy the  $F$ -unique implementation or the  $F$ -unique best-reply properties, we should totally avoid the existence of non-pure equilibrium strategies. For that we define the following property

**Definition 10.** *Given any renegotiation rule  $F \in \mathcal{F}$ , let  $\Gamma = (S, g)$  be a mechanism implementing the FBCR in  $F$ -Nash equilibrium. We say that  $\Gamma$  satisfies the  **$F$ -only pure equilibria property** when, for all  $\theta^{kl} \in \Theta$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma, \theta^{kl})$ ,  $\sigma_A$  and  $\sigma_B$  are pure strategies.*

That is, if a mechanism satisfies this property, all equilibrium strategies are pure. Obviously, for all  $F \in \mathcal{F}$ , the  $F$ -only pure equilibria property implies the  $F$ -pure equilibrium property. As we have already argued, we will always look for mechanisms satisfying the  $F$ -pure equilibrium property, and in case that these mechanisms also satisfy the  $F$ -unique implementation property or the  $F$ -unique best-reply property, then we will also want that they verify the  $F$ -only pure equilibria property.

As usual, we will write (only) pure equilibria property to denote the  $F^I$ -(only) pure equilibrium property. In the same way, from now, for all mechanism  $\Gamma = (S, g)$  we will still write  $s = (s_A, s_B) \in S$  to denote the pure strategy where each agent  $i \in \{A, B\}$  is announcing with probability one  $s_i$ .

To summarize, these are the properties that we consider desirable for our implementing mechanisms: (i) strategy-spaces as small as possible, (ii) enforceability, (iii) renegotiation-proofness, (iv)  $F$ -unique implementation, (v)  $F$ -unique best-reply, and (vi)  $F$ -only pure equilibria. Regarding properties (iii)-(vi), suppose that we are able to design a renegotiation-proof mechanism. Since the basic idea behind renegotiation-proofness is that any mechanism satisfying this property implements the FBCR for all  $F \in \mathcal{F}$ , we would also like that such mechanism satisfies the  $F$ -unique implementation, the  $F$ -unique best-reply, and the  $F$ -only pure equilibria properties, not for some given renegotiation rule, but for all  $F \in \mathcal{F}$ . We will discuss about the connection between these properties in the next section.

## 4. COMPATIBILITY OF DESIRABLE PROPERTIES

Unfortunately, some of the properties analyzed in the previous section are not compatible. First, we will prove that any enforceable mechanism satisfying the pure equilibrium property is not renegotiation-proof. For that we need five previous lemmas:

**Lemma 1.** *There exists some  $F \in \mathcal{F}$  with  $v_B(F_B((c_2^*, c_2^*)), \theta^{12}) > v_B(c_2^*, \theta^{12})$ .*

We omit the proof of this lemma. It follows from Assumptions 1 and 2 and from the definition of  $c_2^*$ . Roughly speaking this lemma says that, if agent  $A$  is of type 1 and agent  $B$  is of type 2, the redistribution of  $2c_2^*$  where both agents receive  $c_2^*$  is not Pareto-efficient. In Figure 4.1 we have represented this redistribution problem by means of the Edgeworth box. There we can see that there are some redistributions (like the one given by point  $z$ ), where both agents are strictly better than with  $c_2^*$ .

The next lemma says that, if we give to the agents the first-best contracts, they will not renegotiate:

**Lemma 2.** *For all  $F \in \mathcal{F}$  and all  $\theta^{kl} \in \Theta$ , we have that  $F((c_k^*, c_l^*), \theta^{kl}) = (c_k^*, c_l^*)$ .*

**Proof.** Suppose by contradiction and w.l.o.g. that there is some  $F \in \mathcal{F}$  and some  $\theta^{kl} \in \Theta$  such that  $F_A((c_k^*, c_l^*), \theta^{kl}) \neq c_k^*$ . Then, since  $v_A(F_A((c_k^*, c_l^*)), \theta^{kl}) \geq v_A(c_k^*, \theta^{kl})$ , by definition of  $c_k^*$  we have that  $\pi_{F_A((c_k^*, c_l^*), \theta^{kl})} < \pi_{c_k^*}$ . In that case  $\pi_{F_B((c_k^*, c_l^*), \theta^{kl})} > \pi_{c_l^*}$ . Then, by definition of  $c_l^*$ ,  $v_B(F_B((c_k^*, c_l^*)), \theta^{kl}) < v_B(c_l^*, \theta^{kl})$ , which contradicts that  $F \in \mathcal{F}$ . ■

The following lemma states that, if a mechanism is renegotiation-proof, then any profile of strategies which is a  $F$ -Nash equilibrium for some  $F \in \mathcal{F}$  is also a Nash equilibrium:

**Lemma 3.** *Let  $\Gamma = (S, g)$  be a renegotiation-proof mechanism. Then, for all  $F \in \mathcal{F}$ , all  $\theta^{kl} \in \Theta$  and all  $s \in F - N(\Gamma, \theta^{kl})$  we have that  $s \in N(\Gamma, \theta^{kl})$ .*

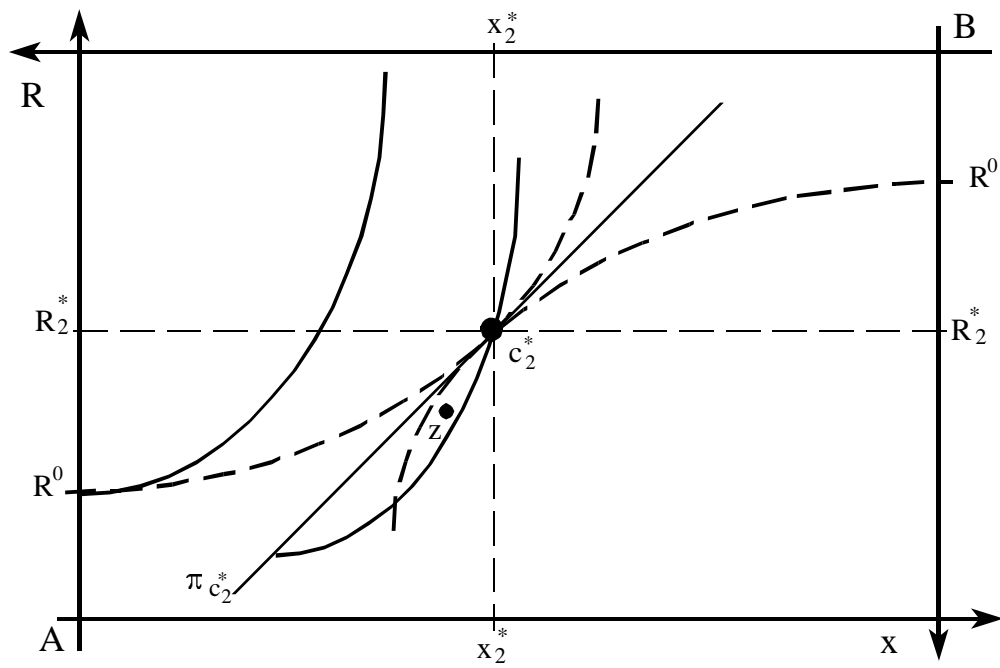


Figure 4.1:



**Proof.** Let  $\Gamma = (S, g)$  be a renegotiation-proof mechanism. Then, for all  $F \in \mathcal{F}$ , all  $\theta^{kl} \in \Theta$  and all  $s \in F - N(\Gamma, \theta^{kl})$  we have that  $g(s) = (c_k^*, c_i^*)$ , and therefore, by Lemma 2,  $F(g(s), \theta^{kl}) = (c_k^*, c_i^*)$ . Suppose by contradiction that for some  $F \in \mathcal{F}$ , some  $\theta^{kl} \in \Theta$  and some  $s \in F - N(\Gamma, \theta^{kl})$  we have that  $s \notin N(\Gamma, \theta^{kl})$ . Then, there exists some  $i \in \{A, B\}$  and some  $s'_i \in S_i$  such that  $v_i(g(s'_i, s_{-i}), \theta^{kl}) > v_i(g(s), \theta^{kl})$ . Therefore, by definition of  $\mathcal{F}$  we have that  $v_i(F(g(s'_i, s_{-i})), \theta^{kl}) \geq v_i(g(s'_i, s_{-i}), \theta^{kl}) > v_i(g(s), \theta^{kl}) = v_i(F(g(s)), \theta^{kl})$ , which contradicts that  $s \in F - N(\Gamma, \theta^{kl})$ . ■

Finally, the next two lemmas show that, in any enforceable mechanism satisfying the pure equilibrium property, all pure Nash equilibrium for the state of the world  $\theta^{12}$  is such that, for some unilateral deviation of agent  $B$ , the mechanism selects contracts  $(c_2^*, c_2^*)$ .

**Lemma 4.** *Let  $\Gamma = (S, g)$  be a mechanism implementing the FBCR in Nash equilibrium and satisfying the pure equilibrium property. For all  $\theta^{kl} \in \Theta$ , let  $(s_A^{kl}, s_B^{kl}) \in N(\Gamma, \theta^{kl})$ . Then we have that  $s_B^{12} \neq s_B^{22}$ .*

**Proof.** Let  $\Gamma = (S, g)$  be a mechanism implementing the FBCR in Nash equilibrium and satisfying the pure equilibrium property. For all  $\theta^{kl} \in \Theta$ , let  $(s_A^{kl}, s_B^{kl}) \in N(\Gamma, \theta^{kl})$ . Suppose by contradiction that  $s_B^{12} = s_B^{22}$ . In that case it is clear that  $s_A^{12} \neq s_A^{22}$  and that  $v(g_A(s_A^{12}, s_B^{22}), \theta^2) = v(g_A(s_A^{12}, s_B^{12}), \theta^2) = v(c_1^*, \theta^2) > v(c_2^*, \theta^2) = v(g_A(s_A^{22}, s_B^{22}), \theta^2)$ , which contradicts that  $(s_A^{22}, s_B^{22}) \in N(\Gamma, \theta^{22})$ . ■

**Lemma 5.** *Let  $\Gamma = (S, g)$  be an enforceable mechanism implementing the FBCR in Nash equilibrium and satisfying the pure equilibrium property. For all  $\theta^{kl} \in \Theta$ , let  $(s_A^{kl}, s_B^{kl}) \in N(\Gamma, \theta^{kl})$ . Then we have that  $g(s_A^{12}, s_B^{22}) = (c_2^*, c_2^*)$ .*

**Proof.** Let  $\Gamma = (S, g)$  be an enforceable mechanism implementing the FBCR in Nash equilibrium and satisfying the pure equilibrium property, and for all  $\theta^{kl} \in \Theta$ , let  $(s_A^{kl}, s_B^{kl}) \in N(\Gamma, \theta^{kl})$ .

Step 1. Since  $(s_A^{22}, s_B^{22}) \in N(\Gamma, \theta^{22})$  it is clear that  $g_A(s_A^{22}, s_B^{22}) = c_2^*$ , and then we have that, for all  $s'_A \in S_A$ , either (i)  $g_A(s'_A, s_B^{22}) \in C$  with  $v(g_A(s'_A, s_B^{22}), \theta^2) = v(c_2^*, \theta^2)$  (and then  $v(g_A(s'_A, s_B^{22}), \theta^1) < U^0$ ), or (ii)  $g_A(s'_A, s_B^{22}) = NC$ .

Step 2. Since  $(s_A^{12}, s_B^{12}) \in N(\Gamma, \theta^{12})$  it is clear that  $g_B(s_A^{12}, s_B^{12}) = c_2^*$ , and then we have that, for all  $s'_B \in S_B$ , either (i)  $g_B(s_A^{12}, s'_B), \theta^2) \in C$  with  $v(g_B(s_A^{12}, s'_B), \theta^2) = v(c_2^*, \theta^2)$  (and then  $v(g_B(s_A^{12}, s'_B), \theta^1) < U^0$ ), or  $g_B(s_A^{12}, s'_B), \theta^2) = NC$ .

Suppose now by contradiction that  $g(s_A^{12}, s_B^{22}) \neq (c_2^*, c_2^*)$ . Then we have four cases:

Case 1. Suppose that  $g(s_A^{12}, s_B^{22}) = (NC, NC)$ . Then by Steps 1 and 2 we have that  $(s_A^{12}, s_B^{22}) \in N(\Gamma, \theta^{11})$ , which contradicts that  $\Gamma$  implements the FBCR in Nash equilibrium.

Case 2. Suppose that  $g_A(s_A^{12}, s_B^{22}) \neq NC$  and  $g_B(s_A^{12}, s_B^{22}) = NC$ . Then by Step 1  $v(g_A(s_A^{12}, s_B^{22}), \theta^2) = v(c_2^*, \theta^2)$ . Therefore, by Steps 1 and 2 we have that  $(s_A^{12}, s_B^{22}) \in N(\Gamma, \theta^{21})$ , which contradicts that  $\Gamma$  implements the FBCR in Nash equilibrium.

Case 3. Suppose that  $g_A(s_A^{12}, s_B^{22}) = NC$  and  $g_B(s_A^{12}, s_B^{22}) \neq NC$ . Then by Step 2  $v(g_B(s_A^{12}, s_B^{22}), \theta^2) = v(c_2^*, \theta^2)$ . Therefore, by Steps 1 and 2 we have that  $(s_A^{12}, s_B^{22}) \in N(\Gamma, \theta^{12})$ , which contradicts that  $\Gamma$  implements the FBCR in Nash equilibrium.

Case 4. Suppose that  $g_A(s_A^{12}, s_B^{22}) \neq NC$  and  $g_B(s_A^{12}, s_B^{22}) \neq NC$ . Then by Steps 1 and 2 we have that  $v(g_A(s_A^{12}, s_B^{22}), \theta^2) = v(c_2^*, \theta^2)$  and  $v(g_B(s_A^{12}, s_B^{22}), \theta^2) = v(c_2^*, \theta^2)$ . Therefore, by Steps 1 and 2 we have that  $(s_A^{12}, s_B^{22}) \in N(\Gamma, \theta^{22})$ . Since we have assumed that  $g(s_A^{12}, s_B^{22}) \neq (c_2^*, c_2^*)$ , this contradicts that  $\Gamma$  implements the FBCR in Nash equilibrium. ■

Now we can prove that enforceability and the pure equilibrium property are not compatible with renegotiation-proofness:

**Proposition 4.** *Let  $\Gamma = (S, g)$  be an enforceable mechanism implementing the FBCR in Nash equilibrium and satisfying the pure equilibrium property. Then  $\Gamma$  is not renegotiation-proof.*

**Proof.** Let  $\Gamma = (S, g)$  be an enforceable mechanism implementing the FBCR in Nash equilibrium. Suppose by contradiction that  $\Gamma$  is renegotiation-proof. Let  $F \in \mathcal{F}$  be such that  $v_B(F_B((c_2^*, c_2^*)), \theta^{12}) > v_B(c_2^*, \theta^{12})$  (by Lemma 1 we know that such  $F$  exists). Let  $s = (s_A, s_B) \in F - N(\Gamma, \theta^{12})$ . Since  $\Gamma$  is renegotiation-proof,  $g(s) = (c_1^*, c_2^*)$ . Then, by Lemma 2  $F_B(g(s), \theta^{12}) = c_2^*$ . Moreover, by Lemma 3  $s \in N(\Gamma, \theta^{12})$ . Then, by Lemmas 4 and 5, there exists some  $s'_B \in S_B$  such that  $g(s_A, s'_B) = (c_2^*, c_2^*)$ . Therefore, we have that  $v_B(F_B(g(s_A, s'_B)), \theta^{12}) = v_B(F_B((c_2^*, c_2^*)), \theta^{12}) > v_B(c_2^*, \theta^{12}) = v_B(F_B(g(s)), \theta^{12})$ , which contradicts that  $s \in F - N(\Gamma, \theta^{12})$ . ■

As we have argued in the previous section, we consider that the pure equilibrium property is a requirement that can not be waived if we want to be sure that the agents will play correctly the mechanism. In that case, from Proposition 4 it is clear that, if we do not know the specific renegotiation rule that will be used by the agents and the domain of those admissible renegotiation rules is  $\mathcal{F}$ , then we will not be able to implement the FBCR by means of an enforceable mechanism.

If we insist on enforceable mechanisms, we will have to impose additional restrictions on the domain of admissible renegotiation rules. That is, we have to be sure that some renegotiation rules in  $\mathcal{F}$  will not be used by the agents (for example, we should be sure that either any renegotiation rule like those defined in Lemma 1 or  $F^I$  will not be used by the agents).

We will say that a subdomain  $\mathcal{F}' \subset \mathcal{F}$  is compatible with enforceability when there exists some enforceable mechanism  $\Gamma = (S, g)$  such that, for all  $F \in \mathcal{F}'$ ,  $\Gamma$  implements the FBCR in  $F$ -Nash equilibrium and it satisfies the  $F$ -pure equilibrium property. For simplicity, when looking for enforceable mechanisms, we will restrict ourselves to the subdomain  $\mathcal{F}' = F^I$  (that is, we will assume that the agents do not renegotiate their contracts). Of course, this is not the only subdomain of  $\mathcal{F}$  compatible with enforceability, but it is the simplest one<sup>10</sup>. Moreover, in that case we will like that the implementing mechanisms only satisfy the unique implementation, unique best-reply and only pure equilibrium properties, instead of requiring the version of these properties for all  $F \in \mathcal{F}$ .

Unfortunately, as the next proposition states, even if we assume that the agents do not renegotiate their contracts among them, enforceability and the pure equilibrium property are not compatible with the unique best-reply property:

**Proposition 5.** *Let  $\Gamma = (S, g)$  be an enforceable mechanism implementing the FBCR in Nash equilibrium and satisfying the pure equilibrium property. Then  $\Gamma$  does not satisfy the unique best-reply property.*

**Proof.** Let  $\Gamma = (S, g)$  be an enforceable mechanism implementing the FBCR in Nash equilibrium and satisfying the pure equilibrium property. For all  $\theta^{kl} \in \Theta$ , let  $(s_A^{kl}, s_B^{kl}) \in N(\Gamma, \theta^{kl})$ . By Lemma 4 we have that  $s_B^{12} \neq s_B^{22}$ . Since  $(s_A^{12}, s_B^{12}) \in N(\Gamma, \theta^{12})$ , we have that  $g_B(s_A^{12}, s_B^{12}) = c_2^*$ . Then, by Lemma 5 we have that  $g_B(s_A^{12}, s_B^{22}) = c_2^* = g_B(s_A^{12}, s_B^{12})$ , which contradicts the unique best-reply property. ■

Finally, although we have no proof of that, we have the intuition that any enforceable mechanism will always have unwanted mixed equilibrium strategies<sup>11</sup>. Then, when we consider enforceable mechanisms, we will implement only in pure strategies.

---

<sup>10</sup>An interesting extension of this work would be to study which is the largest subdomain  $\mathcal{F}' \subset \mathcal{F}$  that is compatible with enforceability.

<sup>11</sup>If a mechanism is enforceable, there is no contract that can be considered as a *bad* outcome for both type of agents. That seems to play a crucial role here.

From these results we have that, regarding the properties stated in the previous section, we will have to choose among the following two desiderata:

**Desiderata 1**

- (1.1) Strategy-spaces as small as possible
- (1.2) Enforceability
- (1.3) Unique implementation

**Desiderata 2**

- (2.1) Strategy-spaces as small as possible
- (2.2) Renegotiation-proofness
- (2.3)  $F$ -unique implementation for all  $F \in \mathcal{F}$
- (2.4)  $F$ -unique best-reply for all  $F \in \mathcal{F}$
- (2.5)  $F$ -only pure equilibria for all  $F \in \mathcal{F}$

## 5. TWO MECHANISMS

In this section we propose two different mechanisms implementing the FBCR in Nash equilibrium and satisfying Desiderata 1 and 2 respectively.

### 5.1. A Mechanism Satisfying Desiderata 1

Before defining the first mechanism we need the following result:

**Lemma 6.** *Let  $c_1^*, c_2^* \in C$  be the first-best contracts for the less and the more productive agents respectively. Then, under Assumptions 1 and 2, there exist some contracts  $c_3, c_4 \in C$  such that:*

- (i)  $v(c_1^*, \theta^1) > v(c_3, \theta^1) > v(c_4, \theta^1) > v(c_2^*, \theta^1)$  and
- (ii)  $v(c_3, \theta^2) > v(c_1^*, \theta^2) > v(c_2^*, \theta^2) = v(c_4, \theta^2)$ .

The proof of this lemma follows from Assumptions 1 and 2 and we omit it. Notice that there are many contracts satisfying Conditions (i) and (ii) of Lemma 6. In Figure 5.1 we provide an example of these contracts. Now, we define Mechanism 1 as follows:

Mechanism 1. Let  $c_3, c_4 \in C$  be as defined in Lemma 6. Then,  $\Gamma_1 = (S, g)$  is such that,  $S_A = \{s_A^1, s_A^2, s_A^3, s_A^4\}$ ,  $S_B = \{s_B^1, s_B^2, s_B^3\}$ , and the outcome function is shown in the following table:

	$s_B^1$	$s_B^2$	$s_B^3$
$s_A^1$	$(c_1^*, c_1^*)$	$(c_2^*, c_3)$	$(c_4, c_2^*)$
$s_A^2$	$(c_2^*, c_4)$	$(c_1^*, c_2^*)$	$(c_2^*, c_2^*)$
$s_A^3$	$(c_3, c_2^*)$	$(c_2^*, c_3)$	$(c_2^*, c_1^*)$
$s_A^4$	$(c_4, c_1^*)$	$(c_3, c_2^*)$	$(c_4, c_2^*)$

Table 1

**Theorem 5.1.** *Mechanism 1 implements the FBCR in Nash equilibrium. Moreover, it satisfies Desiderata 1.*

**Proof.** That the strategy spaces of Mechanism 1 are as small as possible follows from Propositions 1 and 2, and that it is enforceable follows from the fact that for all  $s \in S$  and all  $i \in \{A, B\}$ ,  $v(g_i(s), \theta^2) \geq U^0$ . Now we prove that  $\Gamma_1$  uniquely implements the FBCR in Nash equilibrium.

Claim 1.  $(s_A^1, s_B^1) \in N(\Gamma_1, \theta^{11})$ ,  $(s_A^2, s_B^2) \in N(\Gamma_1, \theta^{12})$ ,  $(s_A^3, s_B^3) \in N(\Gamma_1, \theta^{21})$  and  $(s_A^2, s_B^3) \in N(\Gamma_1, \theta^{22})$ .

On the one hand, for all  $i \in \{A, B\}$  with  $\theta_i = \theta^1$ ,  $c_1^*$  is the best contract that can be selected by  $\Gamma_1$ . On the other hand, for all  $i \in \{A, B\}$  with  $\theta_i = \theta^2$ ,  $c_1^*$  and  $c_3$  are the only contracts which are strictly preferred to  $c_2^*$  for agent  $i$ . Then, from the matrix defining  $\Gamma_1$  it follows that  $(s_A^1, s_B^1) \in N(\Gamma_1, \theta^{11})$ ,  $(s_A^2, s_B^2) \in N(\Gamma_1, \theta^{12})$ ,  $(s_A^3, s_B^3) \in N(\Gamma_1, \theta^{21})$  and  $(s_A^2, s_B^3) \in N(\Gamma_1, \theta^{22})$ .

Claim 2. (2.1) For all  $\theta^{kl} \neq \theta^{11}$ ,  $(s_A^1, s_B^1) \notin N(\Gamma_1, \theta^{kl})$ , (2.2) for all  $\theta^{kl} \neq \theta^{12}$ ,  $(s_A^2, s_B^2) \notin N(\Gamma_1, \theta^{kl})$ , (2.3) for all  $\theta^{kl} \neq \theta^{21}$ ,  $(s_A^3, s_B^3) \notin N(\Gamma_1, \theta^{kl})$  and, (2.4) for all  $\theta^{kl} \neq \theta^{22}$ ,  $(s_A^2, s_B^3) \notin N(\Gamma_1, \theta^{kl})$ .

It follows from the matrix defining  $\Gamma_1$  and the fact that, for all  $i \in \{A, B\}$ ,  $v(c_3, \theta^2) > v(c_1^*, \theta^2)$  and  $v(c_4, \theta^1) > v(c_2^*, \theta^1)$ .

Claim 3. For all  $\theta^{kl} \in \Theta$ , there is no  $s' \in N(\Gamma_1, \theta^{kl})$  with  $s' \notin \{(s_A^1, s_B^1), (s_A^2, s_B^2), (s_A^3, s_B^3), (s_A^2, s_B^3), (s_A^4, s_B^1)\}$ .

Let  $s' = (s'_A, s'_B) \in S$  be such that  $s' \notin \{(s_A^1, s_B^1), (s_A^2, s_B^2), (s_A^3, s_B^3), (s_A^2, s_B^3), (s_A^4, s_B^1)\}$ . Then, there is some  $i \in \{A, B\}$  with  $g_i(s') = c_2^*$  and such that, for some  $s''_i \in S_i$ ,  $g_i(s''_i, s'_{-i}) = c_1^*$ . Since for all  $\theta_i \in \{\theta^1, \theta^2\}$ ,  $v(c_1^*, \theta_i) > v(c_2^*, \theta_i)$ , we have that, for all  $\theta^{kl} \in \Theta$ ,  $s' \notin N(\Gamma_1, \theta^{kl})$ . Moreover, since for all  $\theta_A \in \{\theta^1, \theta^2\}$ ,  $v(c_1^*, \theta_A) > v(c_4, \theta_A)$ , we have that, for all  $\theta^{kl} \in \Theta$ ,  $(s_A^4, s_B^1) \notin N(\Gamma_1, \theta^{kl})$ . ■

Notice that as Proposition 4 predicts, Mechanism 1 is not renegotiation-proof (specifically, for all  $F \in \mathcal{F}$  such that  $v_B(F_B((c_2^*, c_2^*), \theta^{12})) > v_B(c_2^*, \theta^{12})$  and all  $s \in S$  with  $g(s) = (c_1^*, c_2^*)$ , we have that  $s \notin F - N(\Gamma, \theta^{12})$ ), and as Proposition 5

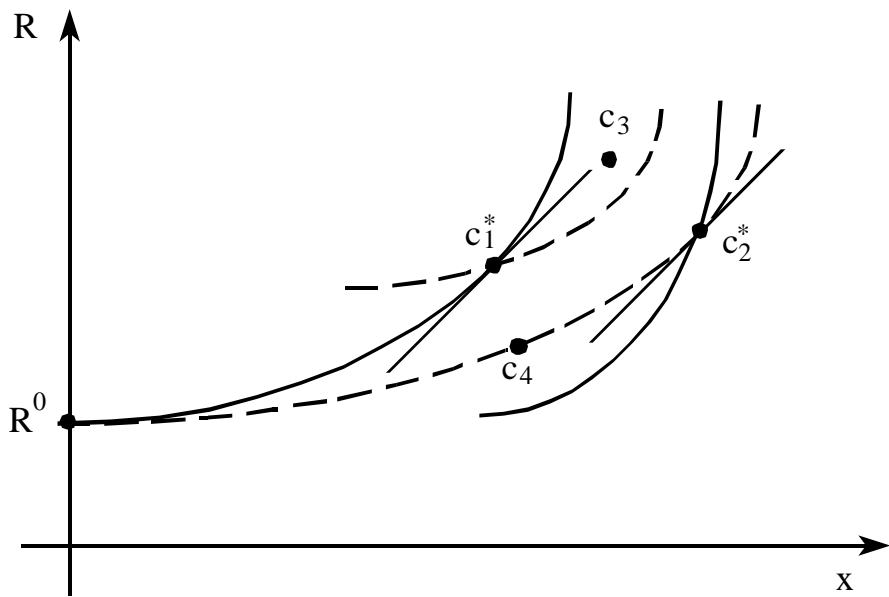


Figure 5.1:

predicts, it does not satisfy the unique best-reply property (for instance,  $(s_A^2, s_B^2) \in N(\Gamma_1, \theta^{12})$ , but  $v_B(g((s_A^2, s_B^2), \theta^{12})) = v_B(g((s_A^2, s_B^3), \theta^{12}))$ ).

## 5.2. A Mechanism Satisfying Desiderata 2

Now we propose a mechanism satisfying Desiderata 2. For all  $c \in C$  and all  $\theta \in \{\theta^1, \theta^2\}$ , let  $\pi(c, \theta) = \max\{\pi_{c'} : v(c', \theta) \geq v(c, \theta)\}$ . Then we have the following result.

**Lemma 7.** *Let  $c_1^*, c_2^* \in C$  be the first-best contracts for the least and the most productive agents respectively. Then, under Assumptions 1 and 2, there exist some contracts  $c_5, c_6 \in C$  where:*

- (1)  $c_5 \in C$  is such that  $\pi_{c_5} = \pi(c_5, \theta^2) = \pi_{c_1^*}$ ,
- (2)  $c_6 \in C$  is such that (i)  $\pi_{c_6} = \pi(c_6, \theta^2)$ , (ii)  $\pi_{c_6} > \max\{\pi(c_2^*, \theta^1), \pi(c_5, \theta^1)\}$ , and  $2\pi_{c_6} > \pi_{c_2^*} + \pi(c_6, \theta^1)$ .

We omit the proof of this lemma. In Figure 5.2 we provide an example of these contracts. Now we can define the mechanism.

**Mechanism 2.** Let  $c_1^*, c_2^*, c_5, c_6 \in C$  be some contracts as defined in Lemma 7. Then,  $\Gamma_2 = (S, g)$ , is such that, for all  $i \in \{A, B\}$ ,  $S_i = \{s_i^{11}, s_i^{12}, s_i^{21}, s_i^{22}\}$ , and the outcome function is the following:

	$s_B^{11}$	$s_B^{12}$	$s_B^{21}$	$s_B^{22}$	$s_B^0$
$s_A^{11}$	$(c_1^*, c_1^*)$	$(NC, c_5)$	$(NC, NC)$	$(NC, NC)$	$(c_1^*, NC)$
$s_A^{12}$	$(NC, NC)$	$(c_1^*, c_2^*)$	$(c_6, c_6)$	$(c_6, c_6)$	$(NC, NC)$
$s_A^{21}$	$(c_5, NC)$	$(NC, NC)$	$(c_2^*, c_1^*)$	$(c_6, c_5)$	$(c_5, NC)$
$s_A^{22}$	$(NC, c_6)$	$(c_5, c_6)$	$(c_6, c_6)$	$(c_2^*, c_2^*)$	$(NC, NC)$
$s_A^0$	$(NC, c_1^*)$	$(NC, c_5)$	$(NC, NC)$	$(NC, NC)$	$(NC, NC)$

Table 2

As we show in the Appendix,  $v(c_6, \theta^2) < U^0$  (and so  $v(c_6, \theta^1) < U^0$ ). Then, it is clear that this mechanism does not satisfy enforceability. Mechanism 2 tries to make up for this lack by including a strategy for each agent  $i \in \{A, B\}$ ,  $s_i^0$ , that can be interpreted as the option of not signing any contract. If an agent decides to play this strategy, he will not sign any contract, regardless of the strategy announced by his counterpart. Although we have had to give up enforceability, at least the mechanism gives to the agents the chance of not signing any contract.

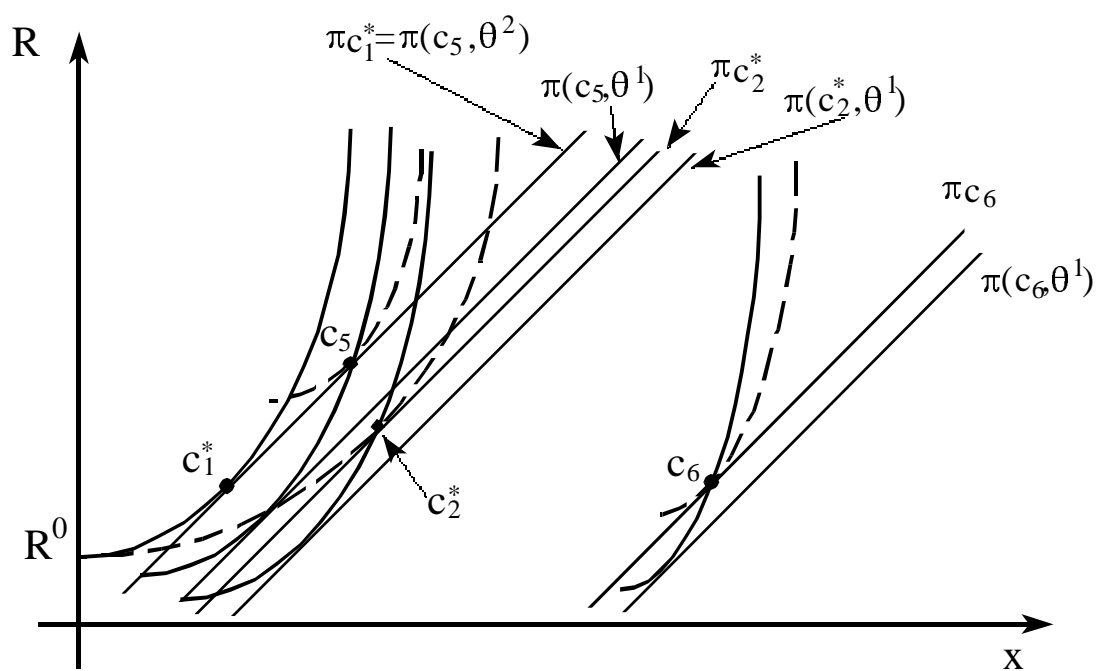


Figure 5.2:



**Theorem 5.2.** *Mechanism 2 satisfies Desiderata 2.*

We provide the proof of this theorem in the Appendix. From Propositions 1 and 2 we have that there are not any other mechanism implementing the FBCR in Nash equilibrium where the strategy spaces of both agents are substantially smaller<sup>12</sup>. The basic idea behind the rest of the proof is the following. First notice that there is a restriction on the minimum utility level that an agent can obtain by renegotiating a profile of contracts: no agent will accept to end up worse after the renegotiation than with the contract that he originally signed. That is, for all state of the world  $\theta^{kl} \in \Theta$ , all profile of contracts  $\bar{c} \in C \times C$  and all agent  $i \in \{A, B\}$ , we have that  $\min\{v_i(F(\bar{c}), \theta^{kl}) : F \in \mathcal{F}\} = v_i(\bar{c}, \theta^{kl})$ . Obviously this also implies a restriction in the maximum utility level that an agent can obtain by renegotiating a profile of contracts (i.e. by renegotiating any given profile of contracts, no agent  $i$  can obtain a utility level as large that he leaves to his counterpart, say agent  $j$ , with a utility level smaller than the one that agent  $j$  would obtain with his initial contract). Taking this into account we prove that, (i) for all state of the world  $\theta^{kl} \in \Theta$ , all renegotiation rule  $F \in \mathcal{F}$ , and all profile of strategies  $(\sigma_A, \sigma_B) \in \Sigma$  different from the pure strategies  $(s_A^{kl}, s_B^{kl})$ ,  $(\sigma_A, \sigma_B)$  is not a  $F$ -Nash equilibrium of Mechanism 2 when the state of the world is  $\theta^{kl}$ , and (ii) for all state of the world  $\theta^{kl} \in \Theta$  and all renegotiation rule  $F \in \mathcal{F}$ , the profile of pure strategies  $(\sigma_A, \sigma_B) = (s_A^{kl}, s_B^{kl})$  is a  $F$ -strict-Nash equilibrium of Mechanism 2 when the state of the world is  $\theta^{kl}$ .

## 6. CONCLUSION

When the agents have complete information the principal can uniquely implement in Nash equilibrium the first-best by means of simple and natural mechanisms that

---

<sup>12</sup>There are another implementing mechanisms where the strategy space of one agent has only three elements and the strategy space of the other agent has four elements (for example Mechanism 1). Obviously, this is not a substantial reduction on the size of the strategy spaces. Moreover, the fact that in Mechanism 2 each agent has five strategies allow us to give an economic interpretation to this strategies. For all  $i \in \{A, B\}$ ,  $s_i^0$  can be interpreted as the option of not signing any contract. Moreover, as it is shown in the proof of Theorem 5.2, we can associate each state of the world with a different strategy for each agent (specifically, for all  $\theta^{kl} \in \Theta$ , the equilibrium strategy of agent  $i \in \{A, B\}$  is  $s_i^{kl}$ ). This is not possible in Mechanism 1: although we can add a new strategies for the agents, it can be proved that in any mechanism satisfying Desiderata 1 there will be always an agent who has the same equilibrium strategy for two different states of the world.

are enforceable in the sense that all the contracts that it selects are individually rational for some type of agent. Moreover, if we renounce to enforceability, we can doubly and uniquely implement the first-best in Nash and strict-Nash equilibrium by means of a simple mechanism that has not mixed strategies. The fact that agents renegotiate among them once they have signed their contracts does not change this result: the same simple and natural mechanism works whatever the renegotiation rule used by the agents is.

There is nothing essential in the definition of our mechanisms that prevent us from extending them to the case where there are an arbitrary number of agents and types. When there are  $n$  agents (where  $n$  is an even integer greater than three) and only two different types, an obvious way to implement first-best contracts is by dividing the agents into groups of two and applying our mechanisms in each of these groups separately.

A possible extension is to environments where the two production processes are complementary. However, here a new problem arises, that of organizational structure (see Baron and Besanko [1]). A priori, it seems that implementation of first-best contracts is more difficult under complementarities since the principal can not threatens with using only one of the two production processes.

## 7. APPENDIX

The proof of Theorem 5.2 is preceded by several lemmas. Throughout all the proof we suppose that Assumptions 1, 2 and 3 are satisfied. The first lemma tell us how are the preferences of the least and the most productive agent over the outcomes selected by Mechanism 2, and how are the principal's profits with some of these outcomes.

**Lemma 8.** *Let  $c_1^*, c_2^*, c_5, c_6 \in C$  be as defined in Lemma 7. Then we have that:*

- (i)  $v(c_1^*, \theta^1) > v(NC, \theta^1) > v(c_5, \theta^1) > v(c_6, \theta^1)$  and  $v(NC, \theta^1) > v(c_2^*, \theta^1) > v(c_6, \theta^1)$ ,
- (ii)  $v(c_5, \theta^2) > v(c_1^*, \theta^2) > v(c_2^*, \theta^2) > v(NC, \theta^2) > v(c_6, \theta^2)$ , and
- (iii)  $\pi_{c_1^*} = \pi_{c_5} < \pi_{c_2^*} < \pi_{c_6}$ .

We omit the proof of this lemma. It follows from former definitions and from Assumptions 1, 2 and 3. We will use these lemma in the rest of the proof, although we will not mention it explicitly. The following lemma states that if both agents are of the same type and they sign the same contract, then no renegotiation is possible.

**Lemma 9.** For all  $\theta^{kk} \in \Theta$  ( $k \in \{1, 2\}$ ), all  $\bar{c} = (c_A, c_B) \in \mathbb{R}_+^{2 \times 2}$  with  $c_A = c_B$ , and all  $F \in \mathcal{F}$ , we have that  $F_A(\bar{c}, \theta^{kk}) = F_B(\bar{c}, \theta^{kk}) = c_A = c_B$ .

**Proof.** Suppose by contradiction that for some  $\theta^{kk} \in \Theta$  ( $k \in \{1, 2\}$ ), some  $\bar{c} = (c_A, c_B) \in \mathbb{R}_+^{2 \times 2}$  with  $c_A = c_B$ , and some  $F \in \mathcal{F}$ , we have that  $F_A(\bar{c}, \theta^{kk}) \neq F_B(\bar{c}, \theta^{kk})$ . Notice that then, for some  $\lambda \in (0, 1)$ ,  $c_A = c_B = \lambda F_A(\bar{c}, \theta^{kk}) + (1 - \lambda)F_B(\bar{c}, \theta^{kk})$ . Suppose w.l.o.g. that  $v_B(F_A(\bar{c}, \theta^{kk})) \geq v_B(F_B(\bar{c}, \theta^{kk}))$ . Since, by assumption, preferences are strictly convex, we have that  $v_B(c_B, \theta^{kk}) > v_B(F_B(\bar{c}, \theta^{kk}))$ , which is a contradiction. Finally, by definition of  $\mathcal{F}$ , it is clear that if  $\bar{c} = (c_A, c_B) \in \mathbb{R}_+^{2 \times 2}$  is such that  $c_A = c_B$ , and  $F \in \mathcal{F}$  is such that  $F_A(\bar{c}, \theta^{kk}) = F_B(\bar{c}, \theta^{kk})$ , then  $F_A(\bar{c}, \theta^{kk}) = F_B(\bar{c}, \theta^{kk}) = c_A = c_B$ . ■

The rest of lemmas establish some limits on the maximum level of utility that the agents can obtain when they renegotiate the profiles of contracts that the mechanism may select.

**Lemma 10.** Let  $c_1^*, c_2^*, c_5, c_6 \in C$  be as defined in Lemma 7. Then we have that:

- (i)  $\max\{v_B(F((c_1^*, c_1^*)), \theta^{12}) : F \in \mathcal{F}\} < v_B(c_5, \theta^{12})$ , and
- (ii)  $\max\{v_A(F((c_1^*, c_1^*)), \theta^{21}) : F \in \mathcal{F}\} < v_A(c_5, \theta^{21})$ .

**Proof.** Suppose by contradiction that, for some  $F \in \mathcal{F}$ ,  $v_B(F((c_1^*, c_1^*)), \theta^{12}) \geq v_B(c_5, \theta^{12})$ . Let  $F((c_1^*, c_1^*), \theta^{12}) = (c'_A, c'_B)$ . Then, by definition of  $c_1^*$  and  $c_5$ , we have that  $\pi_{c'_B} \leq \pi_{c_5} = \pi_{c_1^*}$  and  $c'_B \neq c_1^*$ . Therefore,  $\pi_{c'_A} \geq \pi_{c_1^*}$  and  $c'_A \neq c_1^*$ . Hence, by definition of  $c_1^*$ ,  $v_A(c_1^*, \theta^{12}) > v_A(c'_A, \theta^{12})$ , which contradicts that  $F \in \mathcal{F}$ . The proof of Point (ii) is analogous. ■

**Lemma 11.** Let  $c_1^*, c_2^*, c_5, c_6 \in C$  be the first-best contract for an agent of type 2. Then we have that:

- (i)  $\max\{v_A(F((c_2^*, c_2^*)), \theta^{12}) : F \in \mathcal{F}\} < U^0$ , and
- (ii)  $\max\{v_B(F((c_2^*, c_2^*)), \theta^{21}) : F \in \mathcal{F}\} < U^0$ .

**Proof.** Suppose by contradiction that, for some  $F \in \mathcal{F}$ ,  $v_A(F((c_2^*, c_2^*)), \theta^{12}) \geq U^0$ . Let  $F((c_2^*, c_2^*), \theta^{12}) = (c'_A, c'_B)$ . Then, by definition of  $c_2^*$ , we have that  $\pi_{c'_A} < \pi_{c_2^*}$ , and therefore,  $\pi_{c'_B} > \pi_{c_2^*}$ . Hence, by definition of  $c_2^*$ ,  $v_B(c_2^*, \theta^{12}) > v_B(c'_B, \theta^{12})$ , which contradicts that  $F \in \mathcal{F}$ . The proof of Point (ii) is analogous. ■

**Lemma 12.** Let  $c_1^*, c_2^*, c_5, c_6 \in C$  be as defined in Lemma 7. Then we have that:

- (i)  $\max\{v_A(F((c_2^*, c_1^*)), \theta^{11}) : F \in \mathcal{F}\} < U^0$ ,
- (ii)  $\max\{v_B(F((c_1^*, c_2^*)), \theta^{11}) : F \in \mathcal{F}\} < U^0$ ,

- (iii)  $\max\{v_A(F((c_2^*, c_1^*)), \theta^{12}) : F \in \mathcal{F}\} < U^0$ ,
- (iv)  $\max\{v_B(F((c_1^*, c_2^*)), \theta^{21}) : F \in \mathcal{F}\} < U^0$ ,
- (v)  $\max\{v_A(F((c_1^*, c_2^*)), \theta^{22}) : F \in \mathcal{F}\} < v_A(c_5, \theta^{22})$ , and
- (vi)  $\max\{v_B(F((c_2^*, c_1^*)), \theta^{22}) : F \in \mathcal{F}\} < v_B(c_5, \theta^{22})$ .

**Proof.** Point (i). Suppose by contradiction that, for some  $F \in \mathcal{F}$ ,  $v_A(F((c_2^*, c_1^*)), \theta^{11}) \geq U^0$ . Let  $F((c_2^*, c_1^*), \theta^{11}) = (c'_A, c'_B)$ . Then, by definition of  $c_1^*$  and  $c_2^*$ , we have that  $\pi_{c'_A} \leq \pi_{c_1^*} < \pi_{c_2^*}$ , and therefore  $\pi_{c'_B} > \pi_{c_1^*}$ . Hence, by definition of  $c_1^*$ ,  $v_B(c_1^*, \theta^{11}) > v_B(c'_B, \theta^{11})$ , which contradicts that  $F \in \mathcal{F}$ .

Point (ii). Analogous to Point (i).

Point (iii). Suppose by contradiction that, for some  $F \in \mathcal{F}$ ,  $v_A(F((c_2^*, c_1^*)), \theta^{12}) \geq U^0$ . Let  $F((c_2^*, c_1^*), \theta^{12}) = (c'_A, c'_B)$ . Then, by definition of  $c_1^*$  and  $c_2^*$ , we have that  $\pi_{c'_A} \leq \pi_{c_1^*} < \pi_{c_2^*}$ . Therefore,  $\pi_{c'_B} \geq \pi_{c_2^*}$ . Hence, by definition of  $c_1^*$  and  $c_2^*$ ,  $v_B(c_1^*, \theta^{12}) > v_B(c_2^*, \theta^{12}) \geq v_B(c'_B, \theta^{12})$ , which contradicts that  $F \in \mathcal{F}$ .

Point (iv). Analogous to Point (iii).

Point (v). Suppose by contradiction that, for some  $F \in \mathcal{F}$ ,  $v_A(F((c_1^*, c_2^*)), \theta^{22}) \geq v_A(c_5, \theta^{22})$ . Let  $F((c_1^*, c_2^*), \theta^{22}) = (c'_A, c'_B)$ . Then, by definition of  $c_1^*$ , and  $c_5$ , we have that  $\pi_{c'_A} \leq \pi_{c_5} = \pi_{c_1^*}$  and  $c'_A \neq c_1^*$ . Therefore,  $\pi_{c'_B} \geq \pi_{c_2^*}$  and  $c'_B \neq c_2^*$ . Hence, by definition of  $c_2^*$ ,  $v_B(c_2^*, \theta^{22}) > v_B(c'_B, \theta^{22})$ , which contradicts that  $F \in \mathcal{F}$ .

Point (vi). Analogous to Point (v). ■

**Lemma 13.** Let  $c_1^*, c_2^*, c_5, c_6 \in C$  be as defined in Lemma 7. Then we have that:

- (i)  $\max\{v_A(F((c_6, c_5)), \theta^{11}) : F \in \mathcal{F}\} < U^0$ ,
- (ii)  $\max\{v_B(F((c_5, c_6)), \theta^{11}) : F \in \mathcal{F}\} < U^0$ ,
- (iii)  $\max\{v_A(F((c_6, c_5)), \theta^{12}) : F \in \mathcal{F}\} < U^0$ ,
- (iv)  $\max\{v_B(F((c_5, c_6)), \theta^{21}) : F \in \mathcal{F}\} < U^0$ ,
- (v)  $\max\{v_A(F((c_5, c_6)), \theta^{12}) : F \in \mathcal{F}\} < U^0$ ,
- (vi)  $\max\{v_B(F((c_6, c_5)), \theta^{21}) : F \in \mathcal{F}\} < U^0$ ,
- (vii)  $\max\{v_A(F((c_6, c_5)), \theta^{22}) : F \in \mathcal{F}\} < U^0$ , and
- (viii)  $\max\{v_B(F((c_5, c_6)), \theta^{22}) : F \in \mathcal{F}\} < U^0$ .

**Proof.** Point (i). Suppose by contradiction that, for some  $F \in \mathcal{F}$ ,  $v_A(F((c_6, c_5)), \theta^{11}) \geq U^0$ . Let  $F((c_6, c_5), \theta^{11}) = (c'_A, c'_B)$ . Then, by definition of  $c_1^*$  and  $c_5$ , we have that  $\pi_{c'_A} \leq \pi_{c_1^*} = \pi_{c_5}$ . Therefore,  $\pi_{c'_B} \geq \pi_{c_6} > \pi(c_5, \theta^1)$ . Hence, by definition of  $c_5$  and  $c_6$ ,  $v_B(c_5, \theta^{11}) > v_B(c'_B, \theta^{11})$ , which contradicts that  $F \in \mathcal{F}$ .

Point (ii). Analogous to Point (i).

Point (iii). Suppose by contradiction that, for some  $F \in \mathcal{F}$ ,  $v_A(F((c_6, c_5)), \theta^{12}) \geq U^0$ . Let  $F((c_6, c_5), \theta^{12}) = (c'_A, c'_B)$ . Then, by definition of  $c_1^*$  and  $c_5$ , we have that  $\pi_{c'_A} \leq \pi_{c_1^*} = \pi_{c_5}$ . Therefore,  $\pi_{c'_B} \geq \pi_{c_6}$ . Hence, by definition of  $c_5$  and  $c_6$ ,  $v_B(c_5, \theta^{12}) > v_B(c'_B, \theta^{12})$ , which contradicts that  $F \in \mathcal{F}$ .

Point (iv). Analogous to Point (iii).

Point (v). Suppose by contradiction that, for some  $F \in \mathcal{F}$ ,  $v_A(F((c_5, c_6)), \theta^{12}) \geq U^0$ . Let  $F((c_5, c_6), \theta^{12}) = (c'_A, c'_B)$ . Then, by definition of  $c_1^*$  and  $c_5$ , we have that  $\pi_{c'_A} \leq \pi_{c_1^*} = \pi_{c_5}$  and  $c'_A \neq c_5$ . Therefore,  $\pi_{c'_B} \geq \pi_{c_6}$  and  $c'_B \neq c_6$ . Hence, by definition of  $c_6$ ,  $v_B(c_6, \theta^{12}) > v_B(c'_B, \theta^{12})$ , which contradicts that  $F \in \mathcal{F}$ .

Point (vi). Analogous to Point (v).

Point (vii). Suppose by contradiction that, for some  $F \in \mathcal{F}$ ,  $v_A(F((c_6, c_5)), \theta^{22}) \geq U^0$ . Let  $F((c_6, c_5), \theta^{22}) = (c'_A, c'_B)$ . Then, by definition of  $c_2^*$  and  $c_6$ , we have that  $\pi_{c'_A} \leq \pi_{c_2^*} < \pi_{c_6}$ . Therefore,  $\pi_{c'_B} > \pi_{c_5}$ . Hence, by definition of  $c_5$ ,  $v_B(c_5, \theta^{22}) > v_B(c'_B, \theta^{22})$ , which contradicts that  $F \in \mathcal{F}$ .

Point (viii). Analogous to Point (vii). ■

**Lemma 14.** *Let  $c_1^*, c_2^*, c_5, c_6 \in C$  be as defined in Lemma 7. Then we have that:*

- (i)  $\max\{v_A(F((c_6, c_6)), \theta^{12}) : F \in \mathcal{F}\} < U^0$ ,
- (ii)  $\max\{v_B(F((c_6, c_6)), \theta^{21}) : F \in \mathcal{F}\} < U^0$ ,
- (iii)  $\max\{v_A(F((c_6, c_6)), \theta^{21}) : F \in \mathcal{F}\} < U^0$ , and
- (iv)  $\max\{v_B(F((c_6, c_6)), \theta^{12}) : F \in \mathcal{F}\} < U^0$ .

**Proof.** Point (i). Suppose by contradiction that, for some  $F \in \mathcal{F}$ ,  $v_A(F((c_6, c_6)), \theta^{12}) \geq U^0$ . Let  $F((c_6, c_6), \theta^{12}) = (c'_A, c'_B)$ . Then, by definition of  $c_1^*$  and  $c_6$ , we have that  $\pi_{c'_A} \leq \pi_{c_1^*} < \pi_{c_6}$ , and therefore,  $\pi_{c'_B} > \pi_{c_6}$ . Hence, by definition of  $c_6$ ,  $v_B(c_6, \theta^{12}) \geq v_B(c'_A, \theta^{12})$ , which contradicts that  $F \in \mathcal{F}$ .

Point (ii). Analogous to Point (i).

Point (iii). Suppose by contradiction that, for some  $F \in \mathcal{F}$ ,  $v_A(F((c_6, c_6)), \theta^{21}) \geq U^0$ . Let  $F((c_6, c_6), \theta^{21}) = (c'_A, c'_B)$ . Then, by definition of  $c_2^*$ , we have that  $\pi_{c'_A} \leq \pi_{c_2^*}$ . Therefore, by definition of  $c_2^*$  and  $c_6$ ,  $\pi_{c'_B} > \pi(c_6, \theta^1)$ . Hence,  $v_B(c_6, \theta^{21}) > v_B(c'_B, \theta^{21})$ , which contradicts that  $F \in \mathcal{F}$ .

Point (iv). Analogous to Point (iii). ■

### **Proof of Theorem 5.2.**

First notice that, from the definition of  $\mathcal{F}$ , the minimum level of utility that an agent can obtain after the renegotiation is what he obtained without any

renegotiation. Moreover, whatever the renegotiation rule is, if the mechanism selects  $NC$  for some agent then no renegotiation is possible.

Claim 1. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \neq (s_A^{11}, s_B^{11})$ ,  $(\sigma_A, \sigma_B) \notin F - N(\Gamma_2, \theta^{11})$ .

Step 1.1. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{11})$ ,  $\sigma_B(s_B^{12}) = 0$ .

From the definition of  $\mathcal{F}$  and from Point (ii) of Lemma 12 and Point (ii) of Lemma 13, we have that, for all  $s_A \in S_A$ ,  $\min\{v_B(F(g(s_A, s_B^0))), \theta^{11}\} : F \in \mathcal{F} \geq \max\{v_B(F(g(s_A, s_B^{12}))), \theta^{11}\} : F \in \mathcal{F}$ . Moreover, for all  $s'_A \in S_A$  for which the former inequality is not strict, we have that  $\min\{v_B(F(g(s'_A, s_B^{21}))), \theta^{11}\} : F \in \mathcal{F} > \max\{v_B(F(g(s'_A, s_B^{12}))), \theta^{11}\} : F \in \mathcal{F}$ , and so the result is clear.

Step 1.2. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{11})$ ,  $\sigma_A(s_A^{21}) = 0$ .

From the definition of  $\mathcal{F}$  and from Point (i) of Lemma 12 and Point (i) of Lemma 13 we have that for all  $s_B \in S_B \setminus \{s_B^{12}\}$ ,  $\min\{v_A(F(g(s_A^{11}, s_B))), \theta^{11}\} : F \in \mathcal{F} > \max\{v_A(F(g(s_A^{21}, s_B))), \theta^{11}\} : F \in \mathcal{F}$ . This fact, together with Step 1.1. prove our statement.

Step 1.3. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{11})$ ,  $\sigma_A(s_A^{22}) = 0$ .

From the definition of  $\mathcal{F}$  and from Lemma 9 we have that for all  $s_B \in S_B \setminus \{s_B^{12}\}$ ,  $\min\{v_A(F(g(s_A^{11}, s_B))), \theta^{11}\} : F \in \mathcal{F} > \max\{v_A(F(g(s_A^{22}, s_B))), \theta^{11}\} : F \in \mathcal{F}$ . This fact, together with Step 1.1. prove our statement.

Step 1.4. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{11})$ ,  $\sigma_B(s_B^{22}) = 0$ .

From the definition of  $\mathcal{F}$  and from Lemma 9 we have that for all  $s_A \in S_A \setminus \{s_A^{21}, s_A^{22}\}$ ,  $\min\{v_B(F(g(s_A, s_B^{11}))), \theta^{11}\} : F \in \mathcal{F} > \max\{v_B(F(g(s_A, s_B^{22}))), \theta^{11}\} : F \in \mathcal{F}$ . This fact, together with Steps 1.2 and 1.3 prove our statement.

Step 1.5. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{11})$ ,  $\sigma_B(s_B^{21}) = 0$ .

Identical to Step 1.4.

Step 1.6. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{11})$ ,  $\sigma_A(s_A^{12}) = 0$ .

From the definition of  $\mathcal{F}$  we have that for all  $s_B \in S_B \setminus \{s_B^{12}, s_B^{21}, s_B^{22}\}$ ,  $\min\{v_A(F(g(s_A^{11}, s_B))), \theta^{11}\} : F \in \mathcal{F} > \max\{v_A(F(g(s_A^{12}, s_B))), \theta^{11}\} : F \in \mathcal{F}$ . This fact, together with Steps 1.1, 1.4 and 1.5 prove our statement.

Step 1.7. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{11})$ ,  $\sigma_A(s_A^0) = 0$ .

Identical to Step 1.6.

Step 1.8. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{11})$ ,  $\sigma_B(s_B^0) = 0$ .

From the definition of  $\mathcal{F}$  we have that  $\min\{v_B(F(g(s_A^{11}, s_B^{11}))), \theta^{11}\} : F \in \mathcal{F} > \max\{v_B(F(g(s_A^{11}, s_B^0))), \theta^{11}\} : F \in \mathcal{F}$ . This fact, together with Steps 1.2, 1.3, 1.6 and 1.7 prove our statement.

Claim 2. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \neq (s_A^{12}, s_B^{12})$ ,  $(\sigma_A, \sigma_B) \notin F - N(\Gamma_2, \theta^{12})$ .

Step 2.1. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{12})$ ,  $\sigma_A(s_A^{21}) = 0$ .

From the definition of  $\mathcal{F}$  and from Point (iii) of Lemma 12 and Point (iii) of Lemma 13, we have that, for all  $s_B \in S_B$ ,  $\min\{v_A(F(g(s_A^0, s_B))), \theta^{12}\} : F \in \mathcal{F} \geq \max\{v_A(F(g(s_A^{21}, s_B))), \theta^{12}\} : F \in \mathcal{F}$ . Moreover, for all  $s'_B \in S_B$  for which the former inequality is not strict, we have that  $\min\{v_A(F(g(s_A^{12}, s'_B))), \theta^{12}\} : F \in \mathcal{F} > \max\{v_A(F(g(s_A^{21}, s'_B))), \theta^{12}\} : F \in \mathcal{F}$ , and so the result is clear.

Step 2.2. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{12})$ ,  $\sigma_A(s_A^{22}) = 0$ .

From the definition of  $\mathcal{F}$  and from Point (i) of Lemma 11, Point (v) of Lemma 13 and Point (i) of Lemma 14, we have that for all  $s_B \in S_B$ ,  $\min\{v_A(F(g(s_A^0, s_B))), \theta^{12}\} : F \in \mathcal{F} \geq \max\{v_A(F(g(s_A^{22}, s_B))), \theta^{12}\} : F \in \mathcal{F}$ . Moreover, for all  $s'_B \in S_B$  for which the former inequality is not strict, we have that  $\min\{v_A(F(g(s_A^{11}, s'_B))), \theta^{12}\} : F \in \mathcal{F} > \max\{v_A(F(g(s_A^{22}, s'_B))), \theta^{12}\} : F \in \mathcal{F}$ , and so the result is clear.

Step 2.3. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{12})$ ,  $\sigma_B(s_B^{11}) = 0$ .

From the definition of  $\mathcal{F}$  and from Point (i) of Lemma 10 we have that for all  $s_A \in S_A \setminus \{s_A^{21}, s_A^{22}\}$ ,  $\min\{v_B(F(g(s_A, s_B^{12}))), \theta^{12}\} : F \in \mathcal{F} > \max\{v_B(F(g(s_A, s_B^{11}))), \theta^{12}\} : F \in \mathcal{F}$ . This fact, together with Steps 2.1 and 2.2 prove our statement.

Step 2.4. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{12})$ ,  $\sigma_B(s_B^{21}) = 0$ .

From the definition of  $\mathcal{F}$  and from Point (iv) of Lemma 14 we have that for all  $s_A \in S_A \setminus \{s_A^{21}, s_A^{22}\}$ ,  $\min\{v_B(F(g(s_A, s_B^{12}))), \theta^{12}\} : F \in \mathcal{F} > \max\{v_B(F(g(s_A, s_B^{21}))), \theta^{12}\} : F \in \mathcal{F}$ . This fact, together with Steps 2.1 and 2.2 prove our statement.

Step 2.5. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{12})$ ,  $\sigma_B(s_B^{22}) = 0$ .

Identical to Step 2.4.

Step 2.6. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{12})$ ,  $\sigma_B(s_B^0) = 0$ .

From the definition of  $\mathcal{F}$  we have that for all  $s_A \in S_A \setminus \{s_A^{21}, s_A^{22}\}$ ,  $\min\{v_B(F(g(s_A, s_B^{12}))), \theta^{12}\} : F \in \mathcal{F} > \max\{v_B(F(g(s_A, s_B^0))), \theta^{12}\} : F \in \mathcal{F}$ . This fact, together with Steps 2.1 and 2.2 prove our statement.

Step 2.7. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{12})$ ,  $\sigma_A(s_A^{11}) = 0$ .

From the definition of  $\mathcal{F}$  we have that  $\min\{v_A(F(g(s_A^{12}, s_B^{12}))), \theta^{12}\} : F \in \mathcal{F} > \max\{v_A(F(g(s_A^{11}, s_B^{12}))), \theta^{12}\} : F \in \mathcal{F}$ . This fact, together with Steps 2.3, 2.4, 2.5 and 2.6 prove our statement.

Step 2.8. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{12})$ ,  $\sigma_A(s_A^0) = 0$ .

Identical to Step 2.7.

Claim 3. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \neq (s_A^{21}, s_B^{21})$ ,  $(\sigma_A, \sigma_B) \notin F - N(\Gamma_2, \theta^{21})$ .

Step 3.1. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{21})$ ,  $\sigma_B(s_B^{12}) = 0$ .

From the definition of  $\mathcal{F}$  and from Point (iv) of Lemma 12 and Point (iv) of Lemma 13, we have that, for all  $s_A \in S_A$ ,  $\min\{v_B(F(g(s_A, s_B^0))), \theta^{11}\} : F \in \mathcal{F} \geq \max\{v_B(F(g(s_A, s_B^{12}))), \theta^{11}\} : F \in \mathcal{F}$ . Moreover, for all  $s'_A \in S_A$  for which the former inequality is not strict, we have that  $\min\{v_B(F(g(s'_A, s_B^{21}))), \theta^{11}\} : F \in$

$\mathcal{F}\} > \max\{v_B(F(g(s'_A, s_B^{12})), \theta^{11}) : F \in \mathcal{F}\}$ , and so the result is clear.

Step 3.2. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{21})$ ,  $\sigma_A(s_A^{12}) = 0$ .

From the definition of  $\mathcal{F}$  and from Point (iii) of Lemma 14 we have that for all  $s_B \in S_B \setminus \{s_B^{12}\}$ ,  $\min\{v_A(F(g(s_A^{11}, s_B)), \theta^{21}) : F \in \mathcal{F}\} > \max\{v_A(F(g(s_A^{12}, s_B)), \theta^{21}) : F \in \mathcal{F}\}$ . This fact, together with Step 3.1 prove our statement.

Step 3.3. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{21})$ ,  $\sigma_B(s_B^{22}) = 0$ .

From the definition of  $\mathcal{F}$  and from Point (ii) of Lemma 11 and Point (vi) of Lemma 13, we have that for all  $s_A \in S_A \setminus \{s_A^{12}\}$ ,  $\min\{v_B(F(g(s_A, s_B^0)), \theta^{21}) : F \in \mathcal{F}\} \geq \max\{v_B(F(g(s_A, s_B^{22})), \theta^{21}) : F \in \mathcal{F}\}$ . Moreover, for all  $s'_A \in S_A \setminus \{s_A^{12}\}$  for which the former inequality is not strict, we have that  $\min\{v_B(F(g(s'_A, s_B^{11})), \theta^{21}) : F \in \mathcal{F}\} > \max\{v_B(F(g(s'_A, s_B^{22})), \theta^{21}) : F \in \mathcal{F}\}$ . This fact, together with Step 3.2, prove our statement.

Step 3.4. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{21})$ ,  $\sigma_A(s_A^0) = 0$ .

From the definition of  $\mathcal{F}$  we have that for all  $s_B \in S_B \setminus \{s_B^{12}, s_B^{22}\}$ ,  $\min\{v_A(F(g(s_A^{21}, s_B)), \theta^{21}) : F \in \mathcal{F}\} > \max\{v_A(F(g(s_A^0, s_B)), \theta^{21}) : F \in \mathcal{F}\}$ . This fact, together with Steps 3.1 and 3.3 prove our statement.

Step 3.5. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{21})$ ,  $\sigma_A(s_A^{22}) = 0$ .

From the definition of  $\mathcal{F}$  and from Point (iii) of Lemma 14 we have that for all  $s_B \in S_B \setminus \{s_B^{12}, s_B^{22}\}$ ,  $\min\{v_A(F(g(s_A^{21}, s_B)), \theta^{21}) : F \in \mathcal{F}\} > \max\{v_A(F(g(s_A^{22}, s_B)), \theta^{21}) : F \in \mathcal{F}\}$ . This fact, together with Steps 3.1 and 3.3 prove our statement.

Step 3.6. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{21})$ ,  $\sigma_B(s_B^0) = 0$ .

From the definition of  $\mathcal{F}$  we have that for all  $s_A \in S_A \setminus \{s_A^{12}, s_A^{22}, s_A^0\}$ ,  $\min\{v_B(F(g(s_A, s_B^{11})), \theta^{21}) : F \in \mathcal{F}\} \geq \max\{v_B(F(g(s_A, s_B^0)), \theta^{21}) : F \in \mathcal{F}\}$ . Moreover, for all  $s'_A \in S_A \setminus \{s_A^{12}, s_A^{22}, s_A^0\}$  for which the former inequality is not strict, we have that  $\min\{v_B(F(g(s'_A, s_B^{21})), \theta^{21}) : F \in \mathcal{F}\} > \max\{v_B(F(g(s'_A, s_B^0)), \theta^{21}) : F \in \mathcal{F}\}$ . This fact, together with Steps 3.2, 3.4 and 3.5 prove our statement.

Step 3.7. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{21})$ ,  $\sigma_A(s_A^{11}) = 0$ .

From the definition of  $\mathcal{F}$  and from Point (ii) of Lemma 10 we have that, for all  $s_B \in S_B \setminus \{s_B^{12}, s_B^{22}, s_B^0\}$   $\min\{v_A(F(g(s_A^{21}, s_B)), \theta^{21}) : F \in \mathcal{F}\} > \max\{v_A(F(g(s_A^{11}, s_B)), \theta^{21}) : F \in \mathcal{F}\}$ . This fact, together with Steps 3.1, 3.3, and 3.6 prove our statement.

Step 3.8. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{21})$ ,  $\sigma_B(s_B^{11}) = 0$ .

From the definition of  $\mathcal{F}$  we have that  $\min\{v_B(F(g(s_A^{21}, s_B^{21})), \theta^{21}) : F \in \mathcal{F}\} > \max\{v_B(F(g(s_A^{21}, s_B^{11})), \theta^{21}) : F \in \mathcal{F}\}$ . This fact, together with Steps 3.2, 3.4, 3.5 and 3.7 prove our statement.

Claim 4. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \neq (s_A^{22}, s_B^{22})$ ,  $(\sigma_A, \sigma_B) \notin F - N(\Gamma_2, \theta^{22})$ .

Step 4.1. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{22})$ ,  $\sigma_B(s_B^{21}) = 0$ .



From the definition of  $\mathcal{F}$  and from Lemma 9 and Point (vi) of Lemma 12 we have that for all  $s_A \in S_A$ ,  $\min\{v_B(F(g(s_A, s_B^{22})), \theta^{22}) : F \in \mathcal{F}\} \geq \max\{v_B(F(g(s_A, s_B^{21})), \theta^{22}) : F \in \mathcal{F}\}$ . Moreover, from Lemma 9, for all  $s'_A \in S_A$  for which the former inequality is not strict, we have that  $\min\{v_B(F(g(s'_A, s_B^{12})), \theta^{22}) : F \in \mathcal{F}\} > \max\{v_B(F(g(s'_A, s_B^{21})), \theta^{22}) : F \in \mathcal{F}\}$ , and then our statement is proved.

Step 4.2. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{22})$ ,  $\sigma_A(s_A^0) = 0$ .

From the definition of  $\mathcal{F}$  we have that for all  $s_B \in S_B \setminus \{s_B^{21}\}$ ,  $\min\{v_A(F(g(s_A^{22}, s_B)), \theta^{22}) : F \in \mathcal{F}\} \geq \max\{v_A(F(g(s_A^0, s_B)), \theta^{22}) : F \in \mathcal{F}\}$ . Moreover, for all  $s'_B \in S_B \setminus \{s_B^{21}\}$  for which the former inequality is not strict, we have that  $\min\{v_A(F(g(s_A^{11}, s'_B)), \theta^{22}) : F \in \mathcal{F}\} > \max\{v_A(F(g(s_A^0, s'_B)), \theta^{22}) : F \in \mathcal{F}\}$ . This together with Step 4.1 prove our statement.

Step 4.3. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{22})$ ,  $\sigma_A(s_A^{12}) = 0$ .

From the definition of  $\mathcal{F}$  and from Lemma 9 and Point (v) of Lemma 12 we have that for all  $s_B \in S_B \setminus \{s_B^{21}\}$ ,  $\min\{v_A(F(g(s_A^{22}, s_B)), \theta^{22}) : F \in \mathcal{F}\} \geq \max\{v_A(F(g(s_A^{12}, s_B)), \theta^{22}) : F \in \mathcal{F}\}$ . Moreover, for all  $s'_B \in S_B \setminus \{s_B^{21}\}$  for which the former inequality is not strict, we have that  $\min\{v_A(F(g(s_A^{11}, s'_B)), \theta^{22}) : F \in \mathcal{F}\} > \max\{v_A(F(g(s_A^{12}, s'_B)), \theta^{22}) : F \in \mathcal{F}\}$ . This together with Step 4.1 prove our statement.

Step 4.4. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{22})$ ,  $\sigma_B(s_B^0) = 0$ .

From the definition of  $\mathcal{F}$  we have that for all  $s_A \in S_A \setminus \{s_A^{12}, s_A^0\}$ ,  $\min\{v_B(F(g(s_A, s_B^{22})), \theta^{22}) : F \in \mathcal{F}\} \geq \max\{v_B(F(g(s_A, s_B^0)), \theta^{22}) : F \in \mathcal{F}\}$ . Moreover, for all  $s'_A \in S_A \setminus \{s_A^{12}, s_A^0\}$  for which the former inequality is not strict, we have that  $\min\{v_B(F(g(s'_A, s_B^{12})), \theta^{22}) : F \in \mathcal{F}\} > \max\{v_B(F(g(s'_A, s_B^0)), \theta^{22}) : F \in \mathcal{F}\}$ . This together with Steps 4.2 and 4.3 prove our statement.

Step 4.5. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{22})$ ,  $\sigma_B(s_B^{11}) = 0$ .

From the definition of  $\mathcal{F}$  and from Lemma 9 we have that for all  $s_A \in S_A \setminus \{s_A^{12}, s_A^0\}$ ,  $\min\{v_B(F(g(s_A, s_B^{12})), \theta^{22}) : F \in \mathcal{F}\} \geq \max\{v_B(F(g(s_A, s_B^{11})), \theta^{22}) : F \in \mathcal{F}\}$ . Moreover, for all  $s'_A \in S_A \setminus \{s_A^{12}, s_A^0\}$  for which the former inequality is not strict, we have that  $\min\{v_B(F(g(s'_A, s_B^{22})), \theta^{22}) : F \in \mathcal{F}\} > \max\{v_B(F(g(s'_A, s_B^{11})), \theta^{22}) : F \in \mathcal{F}\}$ . This together with Steps 4.2 and 4.3 prove our statement.

Step 4.6. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{22})$ ,  $\sigma_A(s_A^{11}) = 0$ .

From the definition of  $\mathcal{F}$  we have that for all  $s_B \in S_B \setminus \{s_B^{11}, s_B^{21}, s_B^0\}$ ,  $\min\{v_A(F(g(s_A^{22}, s_B)), \theta^{22}) : F \in \mathcal{F}\} > \max\{v_A(F(g(s_A^{11}, s_B)), \theta^{22}) : F \in \mathcal{F}\}$ . This fact, together with Steps 4.1, 4.4 and 4.5 prove our statement.

Step 4.7. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{22})$ ,  $\sigma_A(s_A^{21}) = 0$ .

From the definition of  $\mathcal{F}$  and from Point (vii) of Lemma 13 we have that for all  $s_B \in S_B \setminus \{s_B^{11}, s_B^{21}, s_B^0\}$ ,  $\min\{v_A(F(g(s_A^{22}, s_B)), \theta^{22}) : F \in \mathcal{F}\} > \max\{v_A(F(g(s_A^{21}, s_B)), \theta^{22}) : F \in \mathcal{F}\}$ .

$s_B)), \theta^{22}) : F \in \mathcal{F}$ . This fact, together with Steps 4.1, 4.4 and 4.5 prove our statement.

Step 4.8. For all  $F \in \mathcal{F}$  and all  $(\sigma_A, \sigma_B) \in F - N(\Gamma_2, \theta^{22})$ ,  $\sigma_B(s_B^{12}) = 0$ .

From the definition of  $\mathcal{F}$  and from Point (viii) of Lemma 13,  $\min\{v_B(F(g(s_A^{22}, s_B^{22}))), \theta^{22}) : F \in \mathcal{F}\} > \max\{v_B(F(g(s_A^{22}, s_B^{12}))), \theta^{22}) : F \in \mathcal{F}\}$ . This fact, together with Steps 4.2, 4.3, 4.6 and 4.7 prove our statement.

Claim 5. For all  $F \in \mathcal{F}$  and all  $\theta^{kl} \in \Theta$ ,  $(s_A^{kl}, s_B^{kl}) \in F - SN(\Gamma_2, \theta^{kl})$ .

Step 5.1. For all  $F \in \mathcal{F}$ ,  $(s_A^{11}, s_B^{11}) \in F - SN(\Gamma_2, \theta^{11})$ .

From the definition of  $\mathcal{F}$  we have that, for all  $s'_A \in S_A \setminus \{s_A^{11}\}$  and all  $s'_B \in S_B \setminus \{s_B^{11}\}$ ,  $\min\{v_A(F(g(s_A^{11}, s_B^{11}))), \theta^{11}) : F \in \mathcal{F}\} > \max\{v_A(F(g(s'_A, s_B^{11}))), \theta^{11}) : F \in \mathcal{F}\}$  and  $\min\{v_B(F(g(s_A^{11}, s_B^{11}))), \theta^{11}) : F \in \mathcal{F}\} > \max\{v_B(F(g(s_A^{11}, s'_B))), \theta^{11}) : F \in \mathcal{F}\}$ , and so our statement is clear.

Step 5.2. For all  $F \in \mathcal{F}$ ,  $(s_A^{12}, s_B^{12}) \in F - SN(\Gamma_2, \theta^{12})$ .

From the definition of  $\mathcal{F}$  and from Point (v) of Lemma 13 we have that, for all  $s'_A \in S_A \setminus \{s_A^{12}\}$ ,  $\min\{v_A(F(g(s_A^{12}, s_B^{12}))), \theta^{12}) : F \in \mathcal{F}\} > \max\{v_A(F(g(s'_A, s_B^{12}))), \theta^{12}) : F \in \mathcal{F}\}$ . In the same way, from the definition of  $\mathcal{F}$  and from Point (ii) of Lemma 14 we have that, for all  $s'_B \in S_B \setminus \{s_B^{12}\}$ ,  $\min\{v_B(F(g(s_A^{12}, s_B^{12}))), \theta^{12}) : F \in \mathcal{F}\} > \max\{v_B(F(g(s_A^{12}, s'_B))), \theta^{12}) : F \in \mathcal{F}\}$ , and so our statement is clear.

Step 5.3. For all  $F \in \mathcal{F}$ ,  $(s_A^{21}, s_B^{21}) \in F - SN(\Gamma_2, \theta^{21})$ .

From the definition of  $\mathcal{F}$  and from Point (iii) of Lemma 14 we have that, for all  $s'_A \in S_A \setminus \{s_A^{21}\}$ ,  $\min\{v_A(F(g(s_A^{21}, s_B^{21}))), \theta^{21}) : F \in \mathcal{F}\} > \max\{v_A(F(g(s'_A, s_B^{21}))), \theta^{21}) : F \in \mathcal{F}\}$ . In the same way, from the definition of  $\mathcal{F}$  and from Point (vi) of Lemma 13 we have that, for all  $s'_B \in S_B \setminus \{s_B^{21}\}$ ,  $\min\{v_B(F(g(s_A^{21}, s_B^{21}))), \theta^{21}) : F \in \mathcal{F}\} > \max\{v_B(F(g(s_A^{21}, s'_B))), \theta^{21}) : F \in \mathcal{F}\}$ , and so our statement is clear.

Step 5.4. For all  $F \in \mathcal{F}$ ,  $(s_A^{22}, s_B^{22}) \in F - SN(\Gamma_2, \theta^{22})$ .

From the definition of  $\mathcal{F}$  and from Lemma 9 and Point (vii) of Lemma 13, we have that, for all  $s'_A \in S_A \setminus \{s_A^{22}\}$ ,  $\min\{v_A(F(g(s_A^{22}, s_B^{22}))), \theta^{22}) : F \in \mathcal{F}\} > \max\{v_A(F(g(s'_A, s_B^{22}))), \theta^{22}) : F \in \mathcal{F}\}$ . In the same way, from the definition of  $\mathcal{F}$  and from Lemma 9 and Point (viii) of Lemma 13, we have that for all  $s'_B \in S_B \setminus \{s_B^{22}\}$ ,  $\min\{v_B(F(g(s_A^{22}, s_B^{22}))), \theta^{22}) : F \in \mathcal{F}\} > \max\{v_B(F(g(s_A^{22}, s'_B))), \theta^{22}) : F \in \mathcal{F}\}$ , and so our statement is clear. ■

## References

- [1] BARON, D. and D. BESANKO, 1992, Information, Control and Organizational Structure, *Journal of Economics and Management Strategy*, 1, 237-275.
- [2] DEMSKY, J. and D. SAPPINGTON, 1984, Optimal Incentive Contracts with Multiple Agents, *Journal of Economic Theory*, 33, 152-171.
- [3] DEMSKY, J. and D. SAPPINGTON, 1983, Multy-Agent Control in Perfectly Correlated Environments, *Economic Letters*, 13, 325-330.
- [4] DUTTA, B. and A. SEN, 1991, A Necessary and Sufficient Condition for Two-Person Nash Implementation, *Review of Economic Studies*, 58, 121-128.
- [5] A. GIBBARD, 1973, Manipulation Voting Schemes: A General Result, *Econometrica*, 41, 587-602.
- [6] J. GLOVER, 1994, A Simpler Mechanism that Stops Agents from Cheating, *Journal of Economic Theory*, 62, 221-229.
- [7] M. O. JACKSON, 1992, Implementation in Undominated Strategies: A Look at Bounded Mechanisms, *Review of Economic Studies*, 59, 757-775.
- [8] M. O. JACKSON and T. PALFREY, Efficiency and Voluntary Implementation in Markets with Repeated Pairwise Bargaining, forthcoming in *Econometrica*.
- [9] MA, C., MOORE, J. and S. TURNBULL, 1988, Stopping Agents from Cheating, *Journal of Economic Theory*, 46, 355-372.
- [10] MACHO, I. and D. PEREZ CASTRILLO, 1994, Introducción a la Economía de la Información, *Ariel Economía*, Barcelona.
- [11] MASKIN, E., 1979, Nash Equilibrium and Welfare Optimality, *mimeo*.

- [12] MOORE, J., 1992, Implementation, Contracts and Renegotiation in Environments with Symmetric Information, *Econometric Society Monographs*, Vol. ESM 20 (Cambridge University Press), 182-282.
- [13] MOORE, J. and R. REPULLO, 1990, Nash Implementation: A Full Characterization, *Econometrica*, 58, 1083-1099.
- [14] REICHELSTEIN, S. and S. REITER, 1988, Game Forms with Minimal Message Spaces, *Econometrica*, 56, 661-692.