Department of Economics- FEA/USP

MODELING COOPERATIVE DECISION SITUATIONS: THE DEVIATION FUNCTION FORM AND THE EQUILIBRIUM CONCEPT

MARILDA SOTOMAYOR

WORKING PAPER SERIES Nº 2013-09

Department of Economics, FEA-USP
Working Paper Nº 2013-09

MODELING COOPERATIVE DECISION SITUATIONS: THE DEVIATION FUNCTION FORM AND THE EQUILIBRIUM CONCEPT

Marilda Sotomayor (marildas@usp.br)

Abstract:

Rosenthal (1972) points out that the coalitional function form may be insufficient to analyze some strategic interactions of the cooperative normal form. His solution consists in representing games in effectiveness form, which explicitly describes the set of possible outcomes that each coalition can enforce by a unilateral deviation from any proposed outcome.

This paper detects some non-appropriateness of the effectiveness representation with respect to the stability of outcomes against coalitional deviations. In order to correct such inadequacies, it is proposed a new model, called deviation function form, which modifies Rosenthal's setting by also modeling the coalition structure, in an appropriate way, and by incorporating new kinds of coalitional interactions, which support the agreements proposed by deviating coalitions. This modification propitiates that the concept of stability of the matching models, viewed as a cooperative equilibrium concept, be translated to any game in the deviation function form and be confronted with the traditional notion of core. Precise answers are given to the questions raised.

Keywords: cooperative equilibrium, core, stability, matching

JEL Codes: C78, D78

MODELING COOPERATIVE DECISION SITUATIONS: THE DEVIATION FUNCTION FORM AND THE EQUILIBRIUM CONCEPT

MARILDA SOTOMAYOR[1]
Universidade de São Paulo
Department of Economics, Cidade Universitária, Av. Prof. Luciano Gualberto 908, 05508-900, São Paulo, SP, Brazil
e-mail: marildas@usp.br

ABSTRACT

Rosenthal (1972) points out that the coalitional function form may be insufficient to analyze some strategic interactions of the cooperative normal form. His solution consists in representing games in effectiveness form, which explicitly describes the set of possible outcomes that each coalition can enforce by a unilateral deviation from any proposed outcome.

This paper detects some non-appropriateness of the effectiveness representation with respect to the stability of outcomes against coalitional deviations. In order to correct such inadequacies, it is proposed a new model, called deviation function form, which modifies Rosenthal's setting by also modeling the coalition structure, in an appropriate way, and by incorporating new kinds of coalitional interactions, which support the agreements proposed by deviating coalitions. This modification propitiates that the concept of stability of the matching models, viewed as a cooperative equilibrium concept, be translated to any game in the deviation function form and be confronted with the traditional notion of core. Precise answers are given to the questions raised.

Key words: cooperative equilibrium, core, stability, matching
JEL: C78, D78

---

[1] This paper was partially supported by CNPq-Brazil and by FIPE. A draft of it was presented in the conference "Roth and Sotomayor: 20 years after", May, 2010, Duke University. It is a revised version of a lecture delivered at the International Conference on Game Theory and Economic Applications at SUNY, Stony Brook, July 2013.

INTRODUCTION

We idealize an environment in which a cooperative decision situation takes place, involving a finite set of agents who can freely communicate and want to form *coalitions.* These agents will be called *players.*

The players involved in a coalition interact among themselves, by acting according to established *rules,* aiming to reach an agreement (or to sign a contract) on the terms that will regulate their participation in the given coalition. An *outcome* is a set of coalitions, whose union is the whole set of players (*coalition structure*), together with the set of agreements reached by the coalitions in any phase of the negotiation process. An outcome is *feasible* if it does not violate the established rules.

One of the features of this cooperative decision situation is that a player might want to enter in more than one coalition, so a coalition structure is not necessarily a partition of the whole set of agents. Also, the agreements reached in a given coalition are independent of the agreements reached in any other coalition. Of course a player derives a utility level in each coalition he[2] enters and has preferences over possible outcomes.

In this context, the natural question is then: *What outcomes can one predict that will occur?*

The answer to this question involves the assumption that the players should take their decisions based on some criterion of rationality, taking into account the consequences of the possible agreements they could make in each coalition they could form. More specifically, we idealize the cooperative decision situation by assuming that all agents are rational and we postulate that the cooperative behavior of the players should be governed by the following line of reasoning:

"Facing a feasible outcome x, any coalition S of players will take any joint action against x (this joint action may involve current partners out of the coalition), whenever such action is allowed by the established rules and all the outcomes that might arise from this particular joint action are preferred to x by all players in S".

[2] For simplicity of exposition, along this paper we will refer to a player as "he".

The consequences of this line of reasoning for the players lead naturally to some kind of equilibrium, which we will call *cooperative equilibrium*. The intuitive idea is that a *feasible outcome* x *is a* ***cooperative equilibrium*** *if there is no coalition whose members can profitably deviate from x,* ***by taking actions that are allowed by the established rules.*** [3]

Since the players take rational decisions and are free to interact coalitionally, we can expect that the outcomes that will occur should be stable against any coalitional deviation. Thus the prediction will be that only cooperative equilibria will occur. [4]

In general terms, the way game theorists use to approach a cooperative decision situation is by constructing a mathematical model. They do that by abstracting from the negotiation process and focusing on what each coalition can obtain, without specifying how to obtain. The model is called *cooperative game.* How much of the details of the rules of the game should be retained is the central issue in the modeling of a cooperative game situation. Certainly, this depends on the purpose of the analysis.

Basically, the actions players can take to play the game are modeled by the set of *feasible outcomes*. However, the feasible outcomes are too general to capture all the relevant details of the rules of the game for the purpose of observing cooperative equilibria. It turns out that no cooperative equilibrium analysis can ignore the set of feasible actions that the players in a coalition are allowed to take in order to deviate from a given feasible outcome. Taking this into account, the game theorists proposed *forms,* which represent special classes of cooperative games, to serve as vehicle for the equilibrium analysis of these games. For example,

---

[3] This concept contrasts with that of core: *An outcome* x *is in the* **core** *if there is no coalition* S *whose members can profitably deviate from* x *by* ***interacting only among themselves***. Therefore, every cooperative equilibrium is a core outcome.

[4] Well known special cases of such a cooperative decision situation are the matching markets. The coalition structure is given by a matching and the individual payoffs of the players only depend on their agreements with their partners. In these markets the intuitive idea of cooperative equilibrium is captured by the concept of ***stability,*** which has been defined locally, for every matching model that has been studied, since Gale and Shapley (1962).

the cooperative normal form represents games in which each player participates in only one coalition and the payoff of a player is conditioned to the actions taken in all coalitions formed. In the characteristic function form, an outcome is represented by the payoffs of the players, so the information with respect to the actions the players take to reach these payoffs is lost. The conditionality that characterizes the payoffs of the players in the cooperative games in the normal form is also lost. Consequently, as observed in Rosenthal (1972), we may have an outcome which is in the core of the game in the characteristic function form but it is not in the core of the game in the normal form.

In an attempt to correct the imperfections of the characteristic function representation, Rosenthal (1972) proposed the **effectiveness form,** which is enough general to model cooperative games in normal form. Then, an outcome might consider the actions which support the payoffs, and the payoffs of a coalition might depend on the actions taken by the players out of the coalition. For a given coalition S and a given outcome x, Rosenthal defined a set of alternative outcome subsets which the members of S can enforce against x.

In the characteristic function form and in the effectiveness form, the joint actions that the members of a coalition can take against a proposed outcome, and that can be captured by these models, are restricted to “interactions among themselves”, so that the cooperative analysis is based on the core. However, when the rules of the game allow the coalitions to do more than “*to merely interact among themselves”*, we may expect that some core outcomes will not occur. (Sotomayor, 1992, 1999, 2010, 2012). In these cases the cooperative equilibrium analysis only based on the core is not the most appropriate approach.

A simple example in the text illustrates that the cooperative equilbrium analysis may be deficient if it uses as vehicle the characteristic function form or the effectiveness form. In this example none of these forms captures all the details of the established rules that are relevant to conclude if a given outcome is or is not a cooperative equilibrium.

It is for solving problems as the one presented in that example that we propose the **deviation function form** (*df*-form, for short). This form is a mathematical model which intends to serve as vehicle for cooperative equilibrium analysis of cooperative decision situations of the type idealized here. Our framework is more general than the effectiveness form and it complements that form by also capturing the kinds of coalitional interactions that support the agreements proposed by deviating coalitions. More specifically, as in the effectiveness form, we also define a function, the **deviation function**, that for each feasible outcome x and coalition S associates **a set of feasible deviations from x via S.** These outcomes intend to reflect, in some sense, which feasible actions the members of S can take against x. Since these feasible actions for the players in S are not necessarily restricted to "interaction only among themselves", the two sets may be distinct and the set of deviation outcomes from x via S contains the set of effective outcomes for S against x.

The intuitive meaning of a feasible deviation from an outcome x via some coalition S leads naturally to the identification of some structure that such deviation should have for capturing the relevant details of the rules of the game for the cooperative equilibrium analysis purpose. Roughly speaking, in a feasible deviation y from x via S, (i) the members of S make new agreements and make these agreements only among them; (ii) if a current coalition of x is maintained in y, and gives to its members higher individual payoffs than those obtained at x, then this current coalition must be contained in S; (iii) any coalition formed with players in S and players out of S must be some current coalition of x; (iv) the interaction inside such coalition must keep the current agreements or reformulate some of the terms of them. As illustrated by an example in the text, the reformulation of the agreements in some current coalition does not increase the individual payoffs of the members of that coalition.

Some internal consistency is required for the set of feasible deviations from an outcome x via some coalition S. If y belongs to this set, then any feasible outcome at which the players in S take the same actions as at y is also a feasible deviation from x via S. Also, if the members of coalition S only interact among themselves, then y is a feasible deviation from x via S. If ,in addition, the payoffs of the players in S only depend on the coalitions they form, then the outcome is a feasible deviation from any feasible outcome via S.

The intuitive idea of cooperative equilibrium for deviation function form games is captured by the solution concept called *stability,* whose definition uses a new kind of domination relation. Roughly speaking, a feasible outcome x is *destabilized* by a coalition S if there is some feasible deviation from x via S such that all the outcomes that arise from this particular deviation are preferred to x by all players in S. An outcome is *stable* if it is not destabilized by any coalition.

For the models where the payoffs of the members of the coalitions only depend on the agreements made inside the coalitions, the definition of stability has a simpler form: *A feasible outcome x is* ***stable*** *for a game in the deviation function form if there is no coalition S and no feasible deviation from x via S which is preferred to x by every player in S.* This definition applies to the matching models, providing a general definition of stability for these models.

The key observation in the modeling of the cooperative decision situation idealized here is that there might exist more than one way to represent an outcome, and some of these representations might lead to incorrect conclusions. For example, consider the outcome at which each of the following pairs of agents, $S_1=\{p_1, q_1\}$, $S_2=\{p_1,q_2\}$ and $S_3=\{p_2,q_3\}$, agrees to work together. Suppose these agreements are independent. Clearly, $C1=\{S_1, S_2, S_3\}$, and $C2=\{S_1 \cup S_2, S_3\}$ can be used to represent the given outcome. Now observe that the alternative outcome C_3 at which p_1 keeps its partnership with q_1, and p_1 and q_3 form a new coalition, is a feasible deviation from the given outcome via $S=\{p_1, q_3\}$. However, this outcome cannot be identified as a feasible deviation from C2 via S, since **$\{p_1,q_1\}$ is not one of the current coalitions of C2.** Thus, the outcome might be **unstable** if it was represented by C1 and might be **stable** if it was represented by C2.

The way we found to avoid such inconsistency in the representation of an outcome was to require that in the modeling of a feasible outcome, the coalitions be *minimal for the respective agreements.* Roughly speaking, a coalition is minimal if its members *cannot reach the* *part of the agreement due to them by rearranging themselves in proper sub-coalitions.* Thus, in the situation above, the given outcome cannot be represented by C2 because $S_1 \cup S_2$ is not minimal. The players in S_1 and in S_2 get the same agreements as they get in $S_1 \cup S_2$, but in two proper sub-coalitions.

In the remaining part of this paper we define the *effectiveness function* for games in which the outcomes are supported by a minimal coalition structure and we use this function to define the core for these games. Then we establish the connection between the core concept and the cooperative equilibrium concept in a game in the *df*-form. We also provide a sufficient condition under which these two concepts are equivalent.

We point out that when the game is derived from the cooperative normal form, there is some connection between the stability concept and the concept of strong equilibrium point[5]: *the stable outcomes at which all players are single are the strong equilibrium points of the strategic game associated.*

Finally we derive the *df*-form of a game given in the characteristic function form, under the assumption that this form fully describes the decision problem in consideration.[6] For such game we prove that the core and the stability concepts are equivalent.

Further details are discussed in the text. In section 2 the cooperative normal form, the coalitional function form and the effectiveness form proposed by Rosenthal (1972) are described and an example is presented with the intent of illustrating some inadequacy of the coalitional function form and the effectiveness form for the cooperative equilibrium analysis purpose. Section 3 is devoted to model a cooperative decision situation in the *df*- form, to explicit the primitives of this model and to propose the axioms that establish the structure of the feasible deviations. Section 4 defines the solution concept of stability for games in the *df*-form as the notion that captures the intuitive idea of cooperative quilibrium for these games. Section 5 defines the core and the

---

[5] The concept of strong equilibrium point is due to Aumann (1967): a profile of strategies such that no coalition can gain by deviating from it while the others retain the same strategies.
[6] These games were called *coalitional games* by Shapley.

effectiveness function for a game in *df*-form. Section 6 is devoted to prove the connection between the core and the stability concepts in games in the *df*-form. Section 7 derives the *df*-form of a coalitional game in the characteristic function form and proves that, in the coalitional games, the stability concept is equivalent to the core concept. Section 8 concludes the paper and section 9 presents some historical remarks on the stability concept.

2. MOTIVATION

The *normal form* of a cooperative game is derived from strategic situations in which agents can gain from cooperation. It consists of (a) a finite set N of players; (b) a strategy set Σ_S associated with each coalition $S \subseteq N$; (c) for each outcome (P,σ), where $P=\{S_1,\ldots,S_k\}$ is a partition of N, and $\sigma=(\sigma_1, \ldots, \sigma_k)$ is a k-tuple of strategies, with $\sigma_j \in \Sigma_{S_j}$, $j=1,\ldots,k$, there is associated an $|N|$-tuple of utility payoffs.[7] Thus, the utility payoff of a player depends on the actions taken in all partition sets belonging to P.

The strategies in Σ_S represent the actions allowed to S by the rules of the game. They involve all members of S and only members of S and are addressed to the members of S, but they may be conditioned to the actions taken by players out of S.

For each non-empty coalition S, the *coalitional function* V specifies a set $V(S) \subseteq R^{|S|}$. Normally, $V(S)$ is interpreted as the set of $|S|$-dimensional payoff-vectors, each of which coalition S can "assure" itself in some sense, through interactions only among its members. A game in coalitional function form is a triple (N,V,H), where H is the set of possible utility outcomes for the players.[8]

The *effectiveness form* of a game was proposed by Rosenthal (1972). A game G in effectiveness form consists of (a) a finite set N of players; (b) a set X of outcomes; (c) an ordinal, vector-valued utility

[7] The difference between this game form and the non-cooperative normal game form is that in the non-cooperative case the partition is always formed with 1-player coalitions.
[8] Additional assumptions are generally required of (N,V,H). The interested reader is referred to Aumann (1967) for a more complete discussion of the coalitional function form. See also Kannai, Y. (1992).

function $u: X \to R^{|N|}$; and (d) for each point $x \in X$, an *effectiveness function,* which maps every coalition $S \subseteq N$ into a collection of subsets of X.[9]

The effectiveness function for any proposed outcome x, should identify, for each coalition S, the set of alternative subsets of outcomes which the members of S can enforce, at least in a first round, against x, by interacting only among themselves.

The core concept for this framework, defined by Rosenthal (1972), is the following:

> *The* ***core of a game in the effectiveness form*** *is defined to be the set of outcomes against which there exists no objection.*

The idea of an objection is very simple. Suppose an outcome x arises in a negotiation process. Suppose that coalition S, through actions that only involve players in S, enforces the set Y. Then S *objects to* x *with objection set* Y *if every point* $y \in Y$ *that might "reasonably" arise is preferred by every member of* S *to* x. *In this case, every such point* y *is called* an *objection* of S against x

Therefore, if S objects to x then, by interacting only among them in a convenient way, the elements of S are able to get higher payoffs than those given at x.

In the coalitional function formulation, it is usually assumed that the actions taken by the players in $N \backslash S$ cannot prevent S from achieving each of the payoff-vectors in $V(S)$. This unconditional aspect of the coalitional function form makes it deficient in capturing certain features of some cooperative decision situations, such as those that can be represented in the normal form.

The effectiveness form representation of a game is intended to correct such deficiency of the coalitional function form. Its main characteristic is that it is adequate to model cooperative games in normal form, since it captures game situations in which the utility levels reached

[9] Originally (d) requires, for each coalition $S \subseteq N$, an *effectiveness function* which maps ever point $x \in X$ into a collection of subsets of X.

by a coalition S also depend on the actions taken by players in N/S. This dependence is not expressed by $V(S)$.[10]

Example 2.1, below, illustrates that there may be some features of the cooperative decision situations that are relevant for the purpose of cooperative equilibrium analysis, which are not modeled either by the effectiveness form or by the coalitional function form. In this example the cooperative analysis based upon the core is not the correct approach.

EXAMPLE 2.1. (The effectiveness form and the characteristic function form do not capture all relevant details of the rules of the game for the purpose of cooperative equilibrium analysis)

Consider a simple market of buying and selling with two sellers, q_1 and q_2, and one buyer p. Let N denote the set of agents. Seller q_1 has 5 units of a good to sell and seller q_2 has 1 unit of the same good. The maximum amount of money buyer p considers to pay for one unit of the good is \$3. This agent has no utility for more than 5 units of the good. The negotiations are made between the buyer and each seller, independently, respecting the quotas of the agents, and the agents identify utility with money. Furthermore, the market allows some kind of flexibility on the number of items negotiated between the buyer and seller q_1: Once the price of one item is negotiated, **the buyer gets a discount of k% over that price if he acquires 5 units of the good.**

Consider the following allocation x at which buyer p acquires 5 units of the good of seller q_1 and pays to him \$1.80 for each unit. Then seller q_1 will get the total payoff of \$9 and the buyer will get the total payoff of \$6. Seller q_2 does not sell anything.

We can see that x is a cooperative equilibrium when the discount is 20% and it is not so when the discount is 10%. Furthermore x is in the core for any k.

In fact, if the discount is 10% and x is proposed, then buyer p and seller q_2 can counter-propose an alternative outcome y that both prefer.

---

[10]The interested reader can see Examples 1 and 2 of Rosenthal (1972) , where an outcome is not in the core of the game represented in normal form but it is in the core of the game in coalitional function form.

At this outcome buyer p reduces, from 5 to 4, the number of units to be acquired from q_1, in order to trade with q_2. Then he pays \$2 for each unit of the good of q_1 and q_2 sells his item to p for \$0.50. We can expect that p and q_2 might want to take these actions because they are allowed by the rules of the market and make both of them better off. The power of p of increasing his payoff is due to q_1's concurs, which is assured by the flexible nature of the agreement with respect to the number of units negotiated. Therefore, ***x*** **cannot be considered a cooperative equilibrium when k=10.**

The point (6,9,0) is the payoff-vector yielded by x, where the first component is the payoff of the buyer, the second component is the payoff of seller q_1 and the third component is the payoff of seller q_2. The outcome y yields the utility payoff (6.5, 8,0.5), which players p and q_2 both prefer.

Now observe that the value of the discount is not informed either by the characteristic function form or by the effectiveness form. By using the characteristic function V one can only conclude that (6.5, 8,0.5) is in $V(N)$ and (6.5,0.5) is not in $V(p,q_2)$, so (6.5, 8,0.5) does not dominate (6,9,0) via coalition $\{p,q_2\}$. Indeed, the payoff-vector (6,9,0) is clearly undominated, so **it is in the core of the coalitional function form of the game.**

Under the effectiveness form it is only possible to know that y is not in any subset of outcomes which can be enforced by $\{p,q_2\}$ against x. Actually, there is no objection against x, so ***x*** **is in the core of the effectiveness form of the game**.

If the discount is 20%, it is easy to verify that there is no way for p to increase his total payoff by only trading with q_2. If p reduces from 5 to 4 the number of units negotiated with q_1, he will have to pay \$2.25 for each unit of the good of q_1. In this case there is no price that can increase the current total payoffs of p and q_2. Also there are no prices that can increase the current total payoffs of the three agents. **Therefore, the outcome *x* is a cooperative equilibrium, so it is in the core, when** k=20.

We can also observe that, for any k, there are no prices that can increase the current total payoffs of the three agents, so x **is in the core for any** k**.**

In sum, it is advantageous for the buyer to acquire all the units of the good of seller q_1 when the discount is greater than or equal to 20%. Otherwise, he will want to also trade with seller q_2.

The point is that a cooperative equilibrium analysis for the market of this example cannot ignore the type of flexibility of the agreements that can be reached. On the other hand, the type of flexibility of the agreements cannot be modeled, either by the effectiveness form or by the characteristic function form. Therefore, **there is no way to conclude from these representations if** x **is or is not a cooperative equilibrium.** ■

This example indicates that a cooperative equilbrium analysis may be deficient if it uses as vehicle the characteristic function form or the effectiveness form. It is for solving problems as the one presented in this example that we propose the **deviation function form** introduced in the next section.

3. MODELING COOPERATIVE GAMES IN THE *df*-FORM

In this session we provide a mathematical model for the cooperative decision situation described in section 1. The modeling is naturally obtained through the selection of the relevant aspects that should be retained for the cooperative equilibrium analysis purpose. The resulting model, which will be called *deviation function form,* intends to correct the deficiencies of the effectiveness form, as those pointed out in Example 2.1. In what follows, if $x \in X_{\mathcal{B}}$ and $B \in \mathcal{B}$, we will denote by x_B the restriction of the outcome x to the coalition B and by x_p the restriction of x to $B=\{p\}$. Given any sets A and B, we will denote by $A \backslash B$ the set of elements that are in A and are not in B.

In such cooperative decision situation there is a set N of participants called *players.* Any subset of N will be called *coalition*. A set of coalitions whose union is N is called *coalition structure.* Each player may enter more than one coalition, so a coalition structure is not necessarily a partition of N. It is assumed that the members of a coalition do not care about who belongs to the other coalitions that their partners might form. We will denote by C the set of feasible coalition structures. These are the coalition structures which do not violate the established rules, which include the restrictions on the number of coalitions a player may form or on the number of units of labor time he owns to distribute among his partners, etc. (In a college admission market, for example, a feasible coalition structure is given by a matching, which respects the quotas of the colleges, and such that no student is admitted in more than one college).

Given a feasible coalition structure β, a set of *feasible agreements* (agreements, for short) ∇_S is associated to each coalition $S\in\beta$. A feasible agreement $\partial_S\in\nabla_S$ models a possible coalitional interaction among the players belonging to S, whether and when S forms, that involves only players in S and that is feasible to be reached if these players interact only among themselves, according to the rules of the game. In general terms, an agreement is feasible for a given coalition if the activities that are specified in this agreement can be feasibly developed by each member of the coalition.

In a feasible coalition structure, two feasible agreements ∂_S and ∂_T are independently reached by S and T, in the sense that the activity developed by S is independent of that developed by T.

An agreement must specify the part of it that is due to each player in S. Denote

$$\nabla\equiv\{\nabla_S;\ S\in C\}.$$

For an abuse of notation we will use ∇_p instead of $\nabla_{\{p\}}$.

Note that, unlike the normal form of a cooperative game, the set of agreements established by a coalition is not, necessarily, restricted to a set of strategies. In the discrete two-sided matching models, for example, a feasible coalition structure is a feasible two-sided matching, which also models the corresponding agreements (who is matched to whom).

Given a coalition structure $\mathcal{B}=\{B_1,B_2,\ldots,B_k\}$, a k-tuple $\partial=(\partial_1,\ldots,\partial_k)$, with $\partial_j\in\nabla_{B_j}$, is called *agreement structure for* $\mathcal{B}$. If ∂ is an agreement structure for $\mathcal{B}$, we say that $\mathcal{B}$ is *compatible* with ∂, and vice-versa.

An outcome is represented by a pair $(\partial;\mathcal{B})$. As discussed in section 1, there might exist more than one way to represent an outcome, and some of these representations might lead to incorrect conclusions about the stability of the outcome we are trying to model. To avoid this problem it is required that, in the modeling of a feasible outcome, the coalitions be *minimal for the respective agreements.* Roughly speaking, a coalition is minimal if its members *cannot reach the part of the agreement due to them by rearranging themselves in proper sub-coalitions.* It is this property that guarantees uniqueness in the representation of an outcome. The resulting representation will be called *agreement configuration.*

Definition 3.1. Let $\mathcal{B}$ *be a coalition structure and let* ∂ be an *agreement structure for* $\mathcal{B}$. *Coalition* $B\in\mathcal{B}$ *is a* ***minimal coalition*** *at* $(\partial,\mathcal{B})$ (*minimal coalition, for short) if its members cannot reach the part of* ∂_B *due to them by rearranging themselves in proper sub-coalitions (not necessarily pairwise disjoint) of* B*. We say that* $\mathcal{B}$ *is a* **minimal coalition structure** *compatible with* ∂ *if every* $B\in\mathcal{B}$ *is a minimal coalition a* $(\partial,\mathcal{B})$.

Definition 3.2. *An* ***agreement configuration*** *is a pair* $(\partial;\mathcal{B})\equiv((\partial_1,B_1)\ldots,(\partial_k,B_k))$, *where* $\mathcal{B}$ *is a minimal coalition structure compatible with* ∂.

For technical convenience we will consider that if some minimal coalition has only one player then this player is reaching an agreement with himself. In this case we say that the player is *single* in this coalition.[11]

If a player p belongs to several minimal coalitions at an agreement configuration, then the agreements reached by p in each minimal coalition are independent. This is the case, for example, of the continuous two-sided matching markets in which the players form multiple partnerships and the utilities are additively separable. (See Sotomayor 1992, 2012).

During the negotiation process that takes place in such environment, we can imagine that a sequence of outcomes will emerge. It is reasonable to expect that in such outcomes no coalition takes actions that are injurious to its own welfare, or that violate the established rules or the restrictions, which the players might have on their participation in the game. E.g., we cannot expect that during the negotiation process a player agrees to get less in some minimal coalition than the minimum he can guarantee himself by playing as a single player; it is reasonable to also expect that if in a labor market, for instance, a firm agrees to hire a set of workers in block, then no subset of them could give it a higher profit by maintaining the current wages of each one, and so on.

In this context, the outcomes that can be formed at the several steps of the negotiation process are the *feasible outcomes*. [12] Thus,

[11] If side payments are not allowed in the one-to-one assignment game of firms and workers of Shapley and Shubik (1972), for example, the minimal coalitions are given by the firm-worker pairs and by the single agents. The agreement configuration is given by a one-to-one feasible matching, which specifies who works to whom, and an agreement structure, which specifies the salary each worker should receive from the firm which hires him. It should be clear that, if side payments are allowed, then a firm-worker pair may not be a minimal coalition at the given outcome. In the well known College Admission model of Gale and Shapley(1962) (with responsive preferences), a minimal coalition is formed by one student and one college, or only by a single student or a single college.

[12] In the college admission model with responsive preferences, for example, every pair formed with a student and a college at a feasible matching is mutually acceptable. In a model where the colleges admit the students in block, a minimal coalition must include all students that are admitted by a college. In this case, at a feasible matching, every college must admit an acceptable group of students, but this group might include individually unacceptable students for the college.

Under the assumption that the players take individual rational decisions and act according to the established rules, the idea of a ***feasible*** *agreement configuration is captured by an outcome that can be formed at some step of the negotiation process. (Note that a feasible outcome is not necessarily a final outcome of the negotiation process). Thus, the feasible outcomes model the actions that the players can take to rationally play the game. The set of feasible outcomes corresponding to a given coalition structure* $\mathcal{B}$ *will be denoted by* $X_{\mathcal{B}}$*; the set of feasible outcomes,* X*, is the union of all* $X_{\mathcal{B}}$*'s.*

The players derive an individual utility level, or individual payoff, in each coalition they enter. That is, there is a function $U_{pB}: X_{\mathcal{B}} \to R$, so that, for each feasible agreement configuration $(\partial;\mathcal{B})$, the utility level enjoyed by player p if he contributes to coalition $B \in \mathcal{B}$ at $(\partial;\mathcal{B})$ is given by $U_{pB}(\partial;\mathcal{B})$. The number $U_{pB}((\partial;\mathcal{B}))$ is called *p's individual payoff at* $(\partial;\mathcal{B})$ *corresponding to coalition* B.[13] Thus, the payoff of a player might be multi-dimensional: if say, p contributes to the minimal coalitions B_1, B_2 and B_3 at $(\partial;\mathcal{B})$, then he gets the array of individual payoffs $\{U_{pB1}((\partial;\mathcal{B})), U_{pB2}((\partial;\mathcal{B})), U_{pB3}((\partial;\mathcal{B}))\}$[14]. The array of individual payoffs of player p is simply called p's *payoff*. The profile of payoffs, one for each player, is called *payoff vector*.

Given an agreement configuration $(\partial;\mathcal{B})$, the coalition structure $\mathcal{B}$, together with the corresponding payoff vector w is called *payoff configuration corresponding to* $(\partial;\mathcal{B})$ and it is denoted by $(w,\mathcal{B})$. We say that w is compatible with $\mathcal{B}$ and vice-versa. The payoff configuration $(w,\mathcal{B})$ (respectively, payoff vector w) is feasible if $(\partial;\mathcal{B})$ is feasible.[15]

---

[13] In some situations, as those represented by a matching market, the value $U_{pB}((\partial;\mathcal{B}))$ only depends on the agreements made by the players in B. In some other situations, as those that can be represented in the cooperative normal form, such value may also depend on the agreements reached in the minimal coalitions that do not contain p.

[14] This contrasts with the characteristic function form and the effectiveness form, in which the payoffs are one-dimensional.

[15] In the Multiple partners assignment game of Sotomayor (1992), for example, the partners must agree on the division of the income they can generate by working together and a player may contribute to more than one partnership. In this model, a feasible

When the agreement structures of the feasible outcomes are given by the payoff vectors of the players, the payoff configurations are the outcomes of the game. (Further discussion related to this subject is provided in section 7).

Of course, the players have preferences over outcomes.

The structure of preferences over the outcomes is modeled by an ***ordinal payoff function u*** *which associates an* $|N|$*-tuple of utility payoffs* $u(x)=(u_1(x),\ldots,u_{|N|}(x))$ *to each outcome* x.

Then, player p prefers the feasible outcome x to the feasible outcome y if $u_p(x)>u_p(y)$; he is indifferent between the two outcomes if $u_p(x)=u_p(y)$.

The several types of coalitional interactions the members of a coalition S are allowed to perform against a feasible outcome x, in order to deviate from x, are modeled by the set $\phi_x(S)$ of *feasible deviations from* x *via* S.

Given a coalition S and an agreement configuration $y=(\partial,\mathcal{B})$, we denote by $\phi^*_x(S,y)$ the set of outcomes that could result if S deviated from x "by taking the same actions it takes at y". It can be interpreted that coalition S is able, through its actions, to deviate from x by restricting the negotiation process to any one of the subsets $\phi^*_x(S,y)$'s. Nevertheless, coalition S is not able to determine the particular outcome in $\phi^*_x(S,y)$ that will result. Clearly, $y\in\phi^*_x(S,y)$, so $\phi^*_x(S,y)\neq\phi$.

The formal definition of $\phi^*_x(S,y)$ requires some preliminaries. Player q is called ***partner*** of S under an agreement configuration $y=(\partial,\mathcal{B})$, if $q\in B$, for some coalition $B\in\mathcal{B}$ such that $B\cap S\neq\phi$. According to this definition, if $S=\{p\}$, the set of partners of $\{p\}$ under $(\partial,\mathcal{B})$ is the union of all coalitions in $\mathcal{B}$ that contain p and so the partners of p may not be concentrated in the same minimal coalition.

---

outcome is given by a feasible payoff configuration, given by a feasible many-to-many matching together with an array of individual payoffs for each player.

Let $y=(\mathcal{O},\mathcal{B})$. We will use the notation $P[S;y]$ to denote the set of all partners of S at y. Therefore,

$$P[S;y]=\{p\in N;\ p\in B \text{ for some } B\in\mathcal{B} \text{ and } B\cap S\neq\phi\}$$

That is, $P[S;y]=\cup B_j$, over all $B_j\in\mathcal{B}$ such that $B_j\cap S\neq\phi$.

Clearly, $S\subseteq P[S;y]$; i.e., each member of S is also partner of S. Set,

$$S^*(y)\equiv\{B_j\in\mathcal{B};\ B_j\subseteq P[S;y]\};\quad \mathcal{O}_{S^*(y)}\equiv\{\mathcal{O}_{Bj}\in\mathcal{O};\ B_j\in S^*(y)\}.$$

That is, $\mathcal{O}_{S^*(y)}$ is the set of agreements reached in the minimal coalitions at y whose intersection with S is non-empty. Define

$$y^{S^*}\equiv(\mathcal{O}_{S^*(y)},S^*(y)) \text{ and } y^{\mathcal{B}\backslash S^*}\equiv(\mathcal{O}\backslash\mathcal{O}_{S^*(y)},\ \mathcal{B}\backslash S^*(y)).$$

That is, y^{S^*} and $y^{\mathcal{B}\backslash S^*}$ are, respectively, the restrictions of y to the set of partners of S and to the set of non-partners of S. Then, for each coalition S, we can decompose y into y^{S^*} and $y^{\mathcal{B}\backslash S^*}$ and we can represent it as $y=(y^{S^*}, y^{\mathcal{B}\backslash S^*})$.

If $x\in X$ and $y\in\phi_x(S)$, the set $\phi^*_x(S,y)$ is then identified with the set of *feasible deviations* $z=(z^{S^*}, z^{\mathcal{B}\backslash S^*})$ *from* x *via* S *such that* $z^{S^*}=y^{S^*}$. That is,

$$\phi^*_x(S,y)\equiv\{z\in\phi_x(S);\ z^{S^*}=y^{S^*}\}.$$

Of course, $\phi_x(S)=\cup\,\phi^*_x(S,y)$, taken over all $y\in\phi_x(S)$.

Faced with x, if coalition S deviates from x according to y^{S^*}, then any outcome of $\phi^*_x(S,y)$ might result. Excepting the case in which $\phi^*_x(S,y)$ is a singleton, the members of S are not able to determine which particular outcome in $\phi^*_x(S,y)$ will arise.

The model we have just described to represent a cooperative decision situation is called **deviation function form.** Here are the primitives of this model:

(a) a ***set*** ***$N = \{1, . . . ,n\}$*** ***of players****;*

(b) a set C *of* ***feasible coalition structures;***

(c) For each coalition structure β *in* C*, a set* X_β *of* ***feasible outcomes compatible with*** β;[16]

(d) for each $p \in N$*, for each coalition structure* β *and* $B \in \beta$*,, with* $p \in B$*, a* ***utility function*** $U_{pB}: X_\beta \to R$*;*

(e) ***an ordinal, vector-valued utility function*** $u: X \to R^n$*, where* X *is the union of the sets* X_β*'s for all coalition structures* β ***;***

(f) for each $x \in X$*, a* ***deviation function*** ϕ_x *from* x*, which maps every coalition* $S \subseteq N$ *into a set of feasible outcomes, called* ***feasible deviations from x via S****.*

Thus, a game in the **deviation function form** is represented by a 6-tuple (N, C, X, U, u, ϕ)**,** where U is the array of utility functions U_{pS}'s and $\phi \equiv \{\phi_x, x \in X\}$**.**

The intuitive meaning of a feasible deviation from an outcome x via some coalition S leads naturally to the identification of some structure that such deviation should have for capturing the relevant details of the rules of the game for the cooperative equilibrium analysis purpose. Roughly speaking, if y is a feasible deviation from x via S, (i) the members of S make new agreements and make these agreements only among them; (ii) if a coalition B' of y is a current coalition B of x and B' contains players in S, and in addition the individual payoffs of the members of B' at y are higher than their current individual payoffs at B, then B must be contained in S; (iii) any coalition formed with players in S and players out of S must be some current coalition of x; (iv) the

[16] Some times, as in the cooperative games derived from the normal form, it might be useful to include the set of agreements for each coalition as one of the primitives of the model. We decided not to do it because the set of feasible outcomes already specifies such set. On the other hand, the set of coalition structures is necessary to make clear that an outcome is supported by a coalition structure.

interaction inside such coalition keeps the current agreements or reformulates some of the terms of them. In this last case, the reformulation of the agreements in some current coalition does not increase the individual payoffs of the members of that coalition. This is illustrated in Example 2.1: if buyer p acquires 5 units of the good of seller q_1 and the negotiated price of one item is \$1.8, then he can reformulate this agreement by reducing from 5 to 4 the number of items negotiated in order to be able to make a transaction with the other seller. If this is done and the discount is of 10%, the seller will increase the price of each unit to \$2. Clearly, the reduction of the number of items will decrease the individual payoffs of the buyer and of seller q_1 in the coalition $\{p,q_1\}$. However, this could be of interest of the buyer if he acquires the item of seller q_2 at a price less than \$1.

Formally we assume the axioms listed below. P1, P2 and P3 establish the structure of the feasible deviations; P4 requires a sort of internal consistence for the set $\phi_x(S)$; P5 and P6 imply that the *df*-form is more general than the effectiveness form and the characteristic function form, respectively.

(P1) *Let* $x \in X_{\mathcal{B}}$ *and* $x' \in X_{\mathcal{B}'}$. *If* $x' \in \phi_x(S)$ then $\forall\ p \in S\ \exists\ B' \in \mathcal{B}'$ *such that* $p \in B'$ *and* $B' \subseteq S$.

That is, if x' is a feasible deviation from x via S, compatible with the coalition structure $\mathcal{B}'$, then every player in S belongs to, at least, one coalition at $\mathcal{B}'$ only formed with players in S.

(P2) *Let* $x=(\partial,\mathcal{B}) \in X_{\mathcal{B}}$ *and* $x'=(\partial',\mathcal{B}') \in X_{\mathcal{B}'}$. *Suppose* $x' \in \phi_x(S)$, $B' \in \mathcal{B}'$, $B' \cap S \neq \phi$ *and* $B'=B$, *for some* $B \in \mathcal{B}$. *If* $U_{pB'}(x') > U_{pB}(x)$ *for all* $p \in B'$ *then* $B' \subseteq S$.

That is, if a new interaction in a current coalition B makes the players in B better off, then all players in B should be in the coalition S.

(P3) *Let* $x=(\partial,\mathcal{B})\in X_{\mathcal{B}}$ *and* $x'=(\partial',\mathcal{B}')\in X_{\mathcal{B}'}$. *If* $x'\in\phi_x(S)$, $B'\in\mathcal{B}'$, $B'\cap S\neq\phi$ *and* $B'\not\subset S$, *then* $B'=B$, *for some* $B\in\mathcal{B}$. *Furthermore,* $\partial_B=\partial'_{B'}$ *or* $U_{pB}(x')\leq U_{pB}(x)$ *for all* $p\in B$.

That is, if a coalition B' of $\mathcal{B}'$ contains elements of S and elements of $N\backslash S$ then B' must be some current coalition of $\mathcal{B}$. Furthermore, no player in B' can get more at x' than he gets at x, for his participation in B'.

(P4) *Let* $x=(\partial,\mathcal{B})\in X_{\mathcal{B}}$. *Let* x' *and* x'' *be feasible outcomes compatible with the coalition structures* $\mathcal{B}'$ *and* $\mathcal{B}''$, *respectively. If* $S^*(x')=S^*(x'')$ *and* x' *agrees with* x'' *on* $S^*(x')$, *then* $x'\in\phi_x(S)$ *if and only if* $x''\in\phi_x(S)$.

(P5) *Let* $x=(\partial,\mathcal{B})\in X_{\mathcal{B}}$. *Let* $x'\in X_{\mathcal{B}'}$, *such that if* $B'\in\mathcal{B}'$, $B'\cap S\neq\phi$ *then* $B'\subseteq S$. Then $x'\in\phi_x(S)$.

That is, all outcomes at which the players in S only interact among themselves are feasible deviations from x via S. P5 implies that every effective outcome for S against x is a feasible deviation from x via S. (Example 2.1 shows that the converse is not always true).

P6 implies that, if in addition, the payoffs of the members of S do not depend on $N\backslash S$ then every feasible deviation from x via S is a feasible deviation from any outcome via S. Therefore, every feasible outcome y, such that the payoff vector corresponding to y^{S^*} belongs *to* $V(S)$, is a feasible deviation from x via S for all feasible outcome x. (Example 2.1 shows that there may be feasible deviations y from x via S such that the payoff vector corresponding to y^{S^*} is not in $V(S)$).

(P6) *Let* $x=(\partial,\mathcal{B})\in X_{\mathcal{B}}$. *Let* $x'\in X_{\mathcal{B}'}$, *such that if* $B'\in\mathcal{B}'$, $B'\cap S\neq\phi$ *then* $B'\subseteq S$. *Furthermore, the utility level reached by a player if he contributes to a coalition* B' *at* x' *only depends on* $x'_{B'}$. *Then* $x'\in\phi_x(S)$ *for all feasible outcome* x.

4. THE SOLUTION CONCEPT OF STABILITY

In this section we define a solution concept that will be called *stability,* formulated for games that can be represented in the *df*-form. This concept captures the intuitive idea of cooperative equilibrium for these games. The definition of stability uses the version of the domination relation introduced below:

Definition 4.1: *Let (N, C, X, U, u, ϕ) be a game in the deviation function form. Let x and y be in X. Outcome y* ***ϕ-dominates*** *outcome x via coalition S if:*

(a) $u_p(y) > u_p(x)$ for all players $p \in S$ and

(b) $y \in \phi_x(S)$.

Since a coalition S cannot determine the particular outcome in $\phi^*_x(S,y)$ that will result from its deviation from x according to y^{S*}, it is assumed that a coalition deviates whenever all its members **are sure** to be better off. Then we can define:

Definition 4.2: *Let (N, C, X, U, u, ϕ) be a game in the deviation function form. The feasible outcome x is* ***destabilized*** *by coalition S if there is some $y \in \phi_x(S)$ such that x is ϕ-dominated by* **every** *outcome in $\phi^*_x(S,y)$ via coalition S. An outcome $x \in X$ is* **stable** *for (N, C, X, U, u, ϕ) if it is not destabilized by any coalition.*

Clearly, the cooperative equilibria for the cooperative games in the *df*-form are the *stable outcomes*.

In some games, as for example the matching models, the individual payoffs of the members of a minimal coalition only depend on the agreements taken inside that coalition. We will denote by G^* the class of games such that, if $y \in \phi_x(S)$, the individual payoffs of the players in S only depend on the part y^{S*} of y. Therefore, the players in S are indifferent between any two outcomes in $\phi^*_x(S,y)$. In the games of this class, if x is ϕ-dominated by y via coalition S, then x is ϕ-dominated via S by every element of $\phi^*_x(S,y)$. Thus, we can write:

Definition 4.3: *An outcome* x *is* ***stable*** *for the game* $(N,C,X,U,u,\phi)\in G^*$ *if it is feasible and it is not* ϕ*-dominated by any feasible outcome via some coalition.*

Therefore, *the restriction of Definition 4.3 to the matching models provides a general definition of stability for these models.*

The following example illustrates these definitions.

EXAMPLE 4.1. Consider a game in the *df*-form derived from the cooperative normal form where $N=\{1,2\}$ and the sets of strategies are given by $\nabla_1=\{\sigma_1, \sigma_2\}$, $\nabla_2=\{\gamma_1, \gamma_2\}$, $\nabla_{12}=\{\sigma_1\gamma_1, \sigma_1\gamma_2, \sigma_2\gamma_1, \sigma_2\gamma_2\}$. A feasible outcome is given by any partition β of N together with any compatible combination of strategies ∂. The payoff function $U((\partial;\beta))$ is given by:

$U(\{\sigma_1\},\{\gamma_1\},\{1\},\{2\})=U(\{\sigma_1\gamma_1\},\{1,2\})=(4,3)$;

$U(\{\sigma_1\},\{\gamma_2\},\{1\},\{2\})=U(\{\sigma_1\gamma_2\},\{1,2\})=(3,4)$;

$U(\{\sigma_2\},\{\gamma_1\},\{1\},\{2\})=U(\{\sigma_2\gamma_1\},\{1,2\})=(2,5)$;

$U(\{\sigma_2\},\{\gamma_2\},\{1\},\{2\})=U(\{\sigma_2\gamma_2\},\{1,2\})=(5,2)$.

Consider $x=(\{\sigma_1\},\{\gamma_1\},\{1\},\{2\})$ and $S=\{2\}$. If we expect that player 1 will choose any of his strategies when player 2 deviates from x, then y and z are in $\phi_x(S)$, where $y=(\{\sigma_1\},\{\gamma_2\},\{1\},\{2\})$ and $z=(\{\sigma_2\},\{\gamma_2\},\{1\},\{2\})$, and $\phi^*_x(S,y)=\phi^*_x(S,z)=\{y,z\}$. We have that $U_2(y)>U_2(x)>U_2(z)$, so x is ϕ-dominated by y via S and is not ϕ-dominated by z via S. Hence, $S=\{2\}$ does not destabilize x. It is a matter of verification that $\{1\}$ does not destabilize y. Also, no deviation from x via $\{1,2\}$ is preferred to x by both players, so $\{1,2\}$ does not destabilize x. Thus, x is stable under this approach.

However, if x is proposed, the players might claim that the demand for stability is too strong. They could rather relax this demand and still gain something from the game. Player 2, for example, might expect that player 1 would keep, at least temporarily, his strategy σ_1 and so y would arise. It then seems reasonable to consider that $\phi^*_x(S,y)=\{y\}$. Under this approach, $S=\{2\}$ destabilizes x and so x is not stable. ■

REMARK 4.1. As discussed in section 1, the main point of the theory developed here can be summarized in the following: *facing a proposed outcome x, a coalition will take a joint action against x whenever such joint action is allowed by the rules of the game and all its members* ***are sure*** *to be better off.* However, as illustrated by Example 4.1, there are cases in which we can expect that such line of reasoning might be relaxed. In general terms, if a game Γ is derived from the cooperative normal form, let us denote by $\Gamma\#$ the strategic game associated to Γ, obtained by identifying, for each player p, the set ∇_p with the set of strategies for player p. Let Ω be the set of outcomes where all players are single. It seems reasonable to base the cooperative behavior of the players upon the following criterion of rationality:

a coalition, which analyses the possibility of deviating from an outcome $x\in\Omega$, takes an ***optimistic view***. (*)

That is, the members of a coalition S will change their current actions whenever they are **sure** to be better off, **at least in a first round**, in which the players out of the coalition still maintain their current agreements. In this context, if S deviates from $x=(\partial,\mathcal{B})$ according to y^{S^*}, then the outcome that results is $(y^{S^*}, x^{\mathcal{B}\setminus S^*})$. Then, $x\in\Omega$ is stable if there is no coalition S whose members can be better off by changing their current actions while the players in N\S do not change theirs. This is precisely how the *strong equilibrium point*, due to Aumann, is defined. We have then established a connection between these two concepts:

Under the assumption (*), the outcome $x=(\partial,\mathcal{B})\in\Omega$ is stable for Γ if and only if the profile of strategies ∂ is a strong equilibrium point for the strategic game $\Gamma\#$. ■

In any case, from the exposed above, it is understood that, *in order to check instabilities we have to know the kinds of deviations which are feasible (and may be expected) to the coalitions of players.* We present below three cases with the intention of merely suggesting the kind of details of the rules of the game that can be captured by the deviation function form. In all of them, (P1)-(P6) are clearly satisfied.

1st case. *The minimal coalition structures compatible with the feasible outcomes are partitions of N.* In this case the players are allowed to enter one minimal coalition at most. It is then implied by (P1) that, given $x \in X$, $S \subseteq N$ and $y \in \phi_x(S)$, there is no minimal coalition at y that contains elements of S and elements of $N \backslash S$. Hence, at all $y \in \phi_x(S)$, all partners of S are in S.

On the other hand, at all $y \in \phi_x(S)$, $y \neq x$, the players of S only perform *standard coalitional interactions against x:* they discard all current agreements at x and make a new agreement $\partial_S \in \nabla_S$, compatible with a feasible set of minimal coalitions whose union is S and whose pairwise intersection is the empty set.

There are two approaches of interest. Under the first one, if y deviates from x via S, the agreements chosen by the players in $N \backslash S$ at y do not affect the utility levels reached at y by the players in S. Thus, the feasible deviations from x via S are independent of x and are given by:

$$\phi(S) \equiv \{ y \in X; S = P[y,S] \}.$$

The other approach is appropriate when the cooperative decision situation is that derived from a strategic game. In such situation, the utility levels reached by the players in S when they act against x also depend on the actions taken by the players that are not partners of S at x. In these cases, the players in S might reasonably expect that their non-partners at x would continue, at least in a first round, to play their part at x. Thus,

$y=(\partial,\mathcal{B})\in X$ is a feasible deviation from $x=(\gamma,\mathcal{D})$ via S

if (i) $S=P[y,S]$ and (ii) if $B_j\in\mathcal{B}$ and $B_j\cap P[S,x]=\phi$, then $B_j=D_k$, for some $D_k\in\mathcal{D}$, and $\partial_j=\gamma_k$.

Remark 4.1 is related to this case. ■

In the following two cases, the players are allowed to enter more than one minimal active coalition and may perform non-standard coalitional interactions. The power of a deviating coalition also depends on interactions among some of its members with some current partners out of the coalition.

2nd case: *The coalition structures associated to the outcomes are not, necessarily, partitions of N and the agreements, in some sense, are flexible.*

The kind of flexibility that is allowed is specified by the rules of the game situation that is being modeled. In Example 2.1, the market allows some kind of flexibility on the number of items negotiated between the buyer and seller q_1: Once the price of one item is negotiated, **the buyer gets a discount of $k\%$ over that price if he acquires 5 units of the good.** Furthermore, the number of units demanded by the buyer is always accepted if this demand can be satisfied by the seller. It might then be of the interest of the buyer to reduce the number of items that was being proposed at x, to be able to trade with the other seller.

Coalitional interactions of this sort are called *agreement reformulations*. They are proposed by members of the deviating coalition. In such coalitional interactions the individual utility levels of the players involved in the coalition do not increase. On the other hand, this kind of agreement allows some of these players (those who are in the deviating coalition) to make new coalitional interactions with players in other coalitions, which might cause some increasing in their total payoffs.

In general, when agreements are flexible, some players in the deviating coalition S may want to keep some of its current partnerships, which contain partners out of S, and to keep or to reformulate some of the agreements of these current partnerships.

Agreement reformulations are not considered new agreements. That is, given $x=(\gamma,\mathcal{D})$ and $D_j \in \mathcal{D}$, a *new agreement* ∂_j for D_j with respect to x is an agreement such that at least one player in D_j prefers ∂_j to γ_j. Furthermore, all players in D_j are in the deviating coalition.

Then, under flexible agreements,

$y=(\partial,\mathcal{B})$ is a feasible deviation from $x=(\gamma,\mathcal{D})$ via S *if for every $(\partial_j,B_j) \in y^{S*}$ either (i)$B_j=D_k$ for some $D_k \in \mathcal{D}$ and $\partial_j=\gamma_k$; or (ii) $B_j=D_k$ for some $D_k \in \mathcal{D}$, with $B_j \not\subset S$ and ∂_j is a reformulation of γ_k; or (iii)$B_j \subseteq S$ and ∂_j is a new agreement. Furthermore, $S=\cup B_k$ over all B_k with $B_k \subseteq S$.*

In Example 2.1, outcome y is a profitable feasible deviation from x via $S=\{p,q_2\}$ when $k=10$. ■

3rd case: *The coalition structures associated to the outcomes are not, necessarily, partitions of N and the agreements, in some sense, are rigid.*

Under a rigid agreement, if some of the terms are altered, then the whole agreement is nullified. In this case, some members of S may want to keep some of its current agreements with partners out of S as in the previous case. However, they are not allowed to reformulate their current agreements. Then it is possible that the members of S arrange themselves into a feasible set of minimal coalitions (i) by discarding some current minimal coalitions of partners (not necessarily all), if needed; (ii) by keeping some others with their respective agreements; and (iii) by forming new sets of partners, with new agreements, only among themselves.

Thus,

y=(∂,𝓑) is a feasible deviation from x=(γ,𝓓) via S *if for every* $(\partial_j,B_j)\in y^{S^*}$ *either (i)* $B_j=D_k$ *for some* $D_k\in\mathcal{D}$ *and* $\partial_j=\gamma_k$*; or (ii)* $B_j\subseteq S$ *and* ∂_j *is a new agreement. Furthermore,* $S=\cup B_k$ *over all* B_k *with* $B_k\subseteq S$*.* ■

5. THE EFFECTIVENESS FUNCTION AND THE CONCEPT OF CORE OF A GAME IN THE *df*-FORM.

In this section we will define the *effectiveness function* for games in which the outcomes are supported by a minimal coalition structure and we will use this function to define the core for these games. As we have discussed in the previous sections, these games can be fully represented in the *df*-form. The *effectiveness function form* of a game in the *df*-form is closely related to the effectiveness form proposed by Rosenthal (1972). Let $\Gamma=(N,C,X,U,u,\phi)$ be a game in the *df*-form. Given a feasible outcome x and a coalition S define

$$E\phi_x(S)\equiv\{y=(\partial,\mathcal{B})\in X;\ y\in\phi_x(S) \text{ and if } B\in\mathcal{B} \text{ then either } B\subseteq S \text{ or } B\subseteq N\backslash S\} \quad \text{(C3)}$$

That is, $E\phi_x(S)$ is the set of feasible deviations from x via S in which the elements of S interact only among themselves. Then if $y\in E\phi_x(S)$ we have that $P[y,S]=S$. The functions $E\phi_x$ are called the *effectiveness functions of* Γ. It must be pointed out that the level of utility reached by a player in S at an outcome $y\in E\phi_x(S)$ may also depend on the agreements made at y by players in $N\backslash S$.

Given the game $\Gamma=(N,C,X,U,u,\phi)$, the 6-tuple $(N,C,X,U,u,E\phi)$, where $E\phi\equiv\{E\phi_x,\ x\in X\}$, is called the *effectiveness function form* of Γ. It is implied by C3 that $E\phi_x(S)\subseteq\phi_x(S)$ for all feasible outcome x and coalition S. However, these two sets may be distinct, as it happens in the situation illustrated by Example 2.1. There, the effectiveness function form is not adequate to represent the cooperative game situation considered. When $E\phi_x(S)=\phi_x(S)$ for every $x \in X$ and every $S \subseteq N$, the effectiveness function form of Γ can be used to fully represent the game Γ. Games that can be fully represented in its effectiveness function form are called *effective games*.

Definition 5.1: *The game* $\Gamma=(N,C,X,U,u,\phi)$ *is an* ***effective game*** *if and only if* $E\phi \equiv \phi$.

If $x \in X$ and $y \in E\phi_x(S)$, the set $\phi^*_x(S,y)$ of feasible deviations from x via S according to y^{S^*} can be identified with a subset of outcomes enforced by S against x, denoted by $E^*\phi_x(S,y)$. Formally,

$$E^*\phi_x(S,y) \equiv \phi^*_x(S,y) = \{z \in E\phi_x(S);\ z^{S^*} = y^{S^*}\}$$

The definition of **core** of a game in the *df*-form (respectively, effectiveness function form) is given by using the domination relation. It is equivalent to the core concept defined by Rosenthal (1972) for a game in the effectiveness form. Formally,

Definition 5.2: *Let* $\Gamma=(N,C,X,U,u,\phi)$ *(respectively,* $\Gamma=(N,C,X,U,u,E\phi)$ *) be a game in the df- form (respectively, effectiveness function form). The feasible outcome* y ***dominates*** *the feasible outcome* x *via coalition* S *if:*

(a) $u_p(y) > u_p(x) \quad \forall p \in S$ *and*

(b) $y \in E\phi_x(S)$.

Definition 5.3: *Let* $\Gamma=(N,C,X,U,u,\phi)$ *(respectively,* $\Gamma=(N,C,X,U,u,E\phi)$ *) be a game in the df- form (respectively, effectiveness function form). The feasible outcome* x *is* ***blocked*** *by coalition* S *if there is some* $y \in E\phi_x(S)$ *such that* x *is dominated by every outcome in* $E^*\phi_x(S,y)$. *An outcome is in the* ***core*** *of the game* Γ *if it is not blocked by any coalition.*

In games belonging to G^*, if x is dominated by some feasible outcome y via some coalition S, then x is dominated by every element of $E^*\phi_x(S,y)$. For these games we can rewrite the definition above:

Definition 5.4: *Let $\Gamma=(N,C,X,U,u,\phi)$ (respectively, $\Gamma=(N,C,X,U,u,E\phi)$) be a game in G^*. An outcome is in the* ***core*** *of Γ if it is feasible and it is not dominated by any feasible outcome via some coalition.*

This is the usual concept of the core defined by the domination relation for games in which the payoff that a coalition S can achieve does not depend on the actions taken by the players in $N\backslash S$.

6. CONNECTION BETWEEN THE STABILITY AND THE CORE CONCEPTS FOR GAMES IN *df*-FORMS.

Theorem 6.1 connects the core concept and the stability concept in a game in the *df*-form: the stability concept can be viewed as a refinement of the core concept. Such connection has been already established in several matching models, where examples show that the core may be smaller than the set of stable outcomes (see Sotomayor 1992, 1999, 2007, 2010 or 2012, for example). This phenomenon is also illustrated in Example 2.1. Theorem 6.2 assures the equivalence between the two concepts for effective games.

Theorem 6.1. *Let $\Gamma=(N,C,X,U,u,\phi)$ be a game in the df-form. Then the set of core outcomes contains the set of stable outcomes.*

Proof. It is immediate from Definitions 5.2, 5.3, 4.1 and 4.2 and the fact that $E\phi_x(S) \subseteq \phi_x(S)$ that if S blocks x then S destabilizes x. ■

In games in which the coalition structures are partitions of N, each player is allowed to enter one minimal coalition at most. Proposition 6.1

implies that, in this case, the effectiveness function form captures all the relevant details of the game for the purpose of cooperative equilibrium analysis.

Proposition 6.1. *Let* $\Gamma=(N,C,X,U,u,\phi)$ *be a game in the df-form. Suppose that every coalition structure in* C *is a partition of* N*. Then* $E\phi\equiv\phi$*. Consequently,* Γ *is an effective game.*

Proof. By C3, we have that $E\phi_x(S)\subseteq\phi_x(S)$ for any feasible outcome x and coalition S. Thus we only have to show the inclusion in the other direction. Since, by hypothesis, the players are allowed to enter one minimal coalition at most, (P1) implies that if $y\in\phi_x(S)$ then every minimal coalition at y is either contained in S or in $N\backslash S$. It is then implied by (C3) that $y\in E\phi_x(S)$. Hence, $\phi_x(S)\subseteq E\phi_x(S)$ and the proof is complete. ■

Theorem 6.2. *Let* $\Gamma=(N,C,X,U,u,\phi)$ *be a game in the df-form. Suppose that* $E\phi\equiv\phi$*. Then, the set of core outcomes equals the set of stable outcomes.*

Proof. Let x be in the core of Γ. Then x is stable, for otherwise Definition 4.2 implies that there is some coalition S and some $y\in\phi_x(S)$ such that x is ϕ-dominated, via S, by every outcome in $\phi^*_x(S,y)$. Since $y\in E\phi_x(S)$, by hypothesis, it follows from Definition 5.2 that x is dominated via S by every outcome in $E\phi^*_x(S,y)$, so Definition 5.3 implies that x is blocked by S, which is a contradiction. The other direction follows from Theorem 6.1. Hence the proof is complete. ■

Corollary 6.1. *Let* Γ *be an effective game. Then, the set of core outcomes equals the set of stable outcomes.*

Proof. It is immediate from Definition 5.1 and Theorem 6.2. ■

Corollary 6.2. *Let* $\Gamma=(N,C,X,U,u,\phi)$ *(respectively,* $\Gamma=(N,C,X,U,u,E\phi)$ *) be a game in the df-form (respectively, effectiveness function form).*

Suppose that every coalition structure in C is a partition of N. Then, the set of core outcomes equals the set of stable outcomes.

Proof. It is immediate from Proposition 6.1 and Theorem 6.2. ■

The converses of Theorem 6.2 and Corollary 6.2 are not true. In the many-to-one assignment game with additively separable utilities (Sotomayor 1992) the core coincides with the set of stable outcomes. Nevertheless, the players of one of the sides may enter more than one minimal coalition. For this model, it is easy to construct examples of an outcome x such that $E\phi_x(S)\neq\phi_x(S)$.

7. COALITIONAL GAMES

The deviation function form is a mathematical model to represent cooperative decision situations in which the object of interest is an agreement configuration. However, it is possible that a game is given *a priori* in the characteristic function form (N,V,H), where an outcome is a payoff-vector of $R^{|N|}$ and the coalition structure is not specified. In this section we will see how our theory applies to such games.

The *df*-form derivation is made under the assumption that the game is a *coalitional game*, i.e., the characteristic function form is a reasonable description of the decision problem in consideration. Then, if the S-vector $v_S\in V(S)$, it can be interpreted that S can take some joint action which yields itself at least v_S. This payoff vector does not depend on the actions taken by non-members of S. It is then convenient to consider that $H=V(N)$, so the coalitional game is described by (N,V) (see, for example, the discussion concerning this assumption in Rosenthal (1972), page 96).

In such a game, it is not specified a coalition structure, but the existence of some coalition structure that can be associated to a given payoff-vector is always guaranteed (the coalition N can always be formed). Since each player receives only one payoff, then any coalition structure compatible to some payoff-vector of $V(N)$ must be a partition of N into pairwise-disjoint coalitions. Thus a feasible outcome is a feasible payoff configuration $(v, \mathcal{B})$ such that $v \in V(N) \subseteq R^{|N|}$ and $\mathcal{B}$ is a partition of N into minimal coalitions.

Let X be the set of the feasible outcomes. Within this context, it can be then interpreted that the actions that the members of a coalition S are allowed to take against a given outcome x are restricted to the interactions among themselves and do not depend on x. Therefore, the *df*-form and the effectiveness function form can be derived from (N,V) by identifying, for every $x \in X$ and $S \in C$,

$$\phi_x(S) = E\phi_x(S) \equiv \{(v,\mathcal{B}) \in X;\ v_S \in V(S)\}$$

The other elements of these forms are naturally specified. Then, let $\Gamma=(N,C,X,U,u,\phi)$ be the *df*-form associated to the coalitional game (N,V). If $y=(w,\beta)$ and $x=(v,\gamma)$ are feasible payoff configurations such that y is a feasible deviation from x via S, we say that w *is a feasible deviation from* v *via* S. Clearly, (v,γ) is in the core of Γ (respectively, stable) if and only if v is in the core (respectively, a stable payoff) of (N,V).

Since every coalition structure is a partition of N, Corollary 6.2 implies that the set of core payoffs equals the set of stable payoffs in $\Gamma=(N,C,X,U,u,\phi)$. Therefore we have proved the following result.

Theorem 7.1. *Let* $\Gamma=(N,C,X,U,u,\phi)$ be the *df*-form associated to the coalitional game (N,V). *Then, in* Γ, *the set of core payoffs equals the set of stable payoffs.*

As it was seen in Example 2.1, Theorem 7.1 may fail to hold when (N,V) is not a coalitional game.

Since $(N,V) \in G^*$ (if the players in S prefer some outcome in $\phi^*_x(S,y)$ to x, then they prefer every outcome in $\phi^*_x(S,y)$ to x), by using Definition 5.4 and Theorem 7.1, the core and the stability definitions can be rewritten as follows:

Definition 7.1: *A feasible payoff-vector w is* ***stable*** *for (respectively, in the* **core** *of) the coalitional game (N,V) if there is no coalition S and no payoff-vector $v \in V(N)$ such that $v_i > w_i$ for all $i \in S$ and $v_S \in V(S)$.*

This is the usual concept of core for coalitional games.

8. CONCLUSION

The idea of focusing on the stability concept rather than on the core concept has been widely explored in the literature of matching markets, since Gale and Shapley (1962). Following the approach of these authors, some attempt has been done in the mathematical modeling of these markets, in the sense of establishing the concept of stability as the one which captures the intuitive idea of equilibrium for the market in consideration: *an outcome is stable if it is not up set by any coalition.* This idea of equilibrium for matching markets is identified with the idea of cooperative equilibrium when these markets are mathematically modeled as cooperative games.[17]

It turns out that the concept of stable allocation has been locally defined for each matching model that has been studied, and it has not always been associated, and it has not always been correctly associated, to the idea of cooperative equilibrium. In the past literature, some confusion has been due to an incorrect definition of stability in Roth (1984). In the recent literature, the term *stable outcome* has been used, some times, in new models, without any justification, especially among some applied specialists who very rarely pose questions regarding the appropriateness of the solution concept they use. The author simply imposes that the outcomes with certain mathematical properties will be called “stable”. The intuition behind the definition is not discussed.

This work grew out of the attempt to give a precise definition of stability for the matching markets and to extend this definition to more

[17] See, for example, Roth and Sotomayor (1990), chapter 8, where the Assignment game of Shapley and Shubik (1971) is treated as a matching market of buyers and sellers and also modeled as a cooperative game in the coalitional function form.

general games. In the matching models, the definition of stability is supported by the matching structure of an outcome. The generalization of such structure to other games was obtained here by using a coalition structure with *minimal* coalitions. If the coalitions in a coalition structure are not minimal, we may have more than one representation of the same outcome. This does not affect the core stability of the outcome, but we may be led to conclude that it is a cooperative equilibrium under one of the representations and it is not so under some other representation. When the coalitions are minimal the outcome has only one representation. Furthermore, the modeling of an outcome by using a minimal coalition structure permits to well define the sets of feasible deviations, crucial for the identification of the stable outcomes with the cooperative equilibria.

Guided by a key example presented here we identified the line of reasoning which supports the cooperative behavior of the members of a *deviating coalition*. [18] Then, the intuitive idea of cooperative equilibrium emerged naturally.

The next step was to define stability as a solution concept that captured such idea. Example 2.1 showed that some aspects of certain cooperative game situations may fail to be captured in the characteristic function form and in the effectiveness form, yielding incorrect conclusions in both game forms. The necessary ingredients to formulate mathematically the stability concept were provided here by the *deviation function form.* This model establishes, for each feasible outcome x and coalition S, the set of *feasible deviations from* x *via* S.

The coalitional interactions illustrated by our examples allowed the identification of the structure that the outcomes should have to capture, among the actions that the members of S can take against x, those which are feasible and are relevant for the cooperative equilibrium analysis purpose. This structure allows to regard an effective outcome for S against x as a feasible deviation from x via S. This way, the *df*-form is adequate to represent games that can be described in the effectiveness form.

[18] Illustrative examples are given in Sotomayor (1992 and 1999).

The *df*-form intends to be adequate to represent a variety of cooperative decision situations, in which the main activity of the agents is to form coalitions. The matching markets configures as the special class of these situations where the coalitional interactions are restricted to pairwise interactions. In the literature there are several concepts of stability for these models: the *corewise-stability*; the *pairwise stability*, due to Gale and Shapley (1962); the *setwise-stability*, defined in Sotomayor (1999), and the *strong stability* defined by the first time in Sotomayor (1992). These concepts capture the idea of cooperative equilibrium in some models and fail to do that in some other models. None of them provides a general definition of stability. The definition presented in this paper fills this gap in the literature of matching models.

9. HISTORIACAL REMARKS ON THE STABILITY CONCEPT

Gale and Shapley (1962) defined the concept of pairwise-stability for the marriage model and for the college admission model, with the objective of identifying the stable matchings in these models, which turn out to be the core points and the strong core points, respectively. The pairwise-stability concept was translated to the continuous matching models and to the other discrete matching models. Sotomayor 1992 introduced the continuous multiple partners assignment game with additively separable utilities. This paper was the first to point out that the corewise-stability is not equivalent to the stability concept. In this model, the set of stable outcomes is characterized as the set of pairwise-stable outcomes, and it is a subset of the core, but this set may be smaller than the core. Following the approach of Roth (1985) for a discrete many-to-many matching model with substitutable preferences, Blair (1988) observed that the core and the set of pairwise-stable matchings, regarded inappropriately as the stability concept by Roth, might be disjoint.

Later, Sotomayor 1999 showed that pairwise stability is inadequate to define stability in the discrete many-to-many matching models, which include Roth's model, and then proposed the concept of setwise-stability for these models.

The strong stability concept was defined in Sotomayor 1992 for the multiple partners assignment game and proved to be equivalent to the pairwise-stability concept. In that paper, the stable outcomes are identified with the set of strong stable outcomes.

This concept was shown to be adequate to define stability for the time-sharing assignment game with flexible agreements and to be inadequate for the time-sharing assignment game with rigid agreements, both models introduced in Sotomayor 2012. This paper introduced the continuous version of setwise-stability and also proved that this concept is not equivalent to the strong stability concept. Setwise-stability is the stability concept for the time-sharing assignment game with rigid agreements and it is not so for the time-sharing assignment game with flexible agreements. The extension of the setwise-stability concept to a non-matching game was presented, by the fist time, in Sotomayor 2010.

An important class of games that has been very little explored in the literature is that of the discrete matching models with indifferences. For these models, a natural solution concept called *Pareto-stability* was introduced in Sotomayor (2011). This paper argues that when a stable matching is weakly blocked by the grand coalition, the players are induced to renegotiate their agreements through acceptable exchange of partners. The resulting matching is still stable and it does not admit any weak Pareto-improvement. This suggests that not all stable outcomes can be associated to the outcomes that we can predict that will occur. The stability concept defined here can still be refined, by requiring the additional property of Pareto-optimality. When preferences are strict, the stability concept and the Pareto-stability concept are equivalent because there is no weak blocking coalition. With indifferences the set of Pareto-stable outcomes may be a proper subset of the set of stable outcomes.

REFERENCES

Aumann, R. (1967) “A survey of cooperative games without side payments”, M. Shubik, ed. *Essays in Math. Economics,* Princeton University Press, 3-22.

Blair, C., (1988) “The lattice structure of the set of stable matchings with multiple partners”, *Mathematics of Operation Research,* vol. 13, 619–628.

Gale, D. and L. Shapley (1962) “College admissions and the stability of marriage”, *American Mathematical Monthly,* 69, 9-15.

Kannai, Y. (1992) “The core and balancedness”, published in the *Handbook of Game Theory with Economic Applications*, edited by Aumann, R. and S. Hart, Elsevier Science Publishers B. V.

Rosenthal, R. (1972) “Cooperative games in effectiveness form”, *Journal of Economic theory*, 5, 88-101.

Roth, A. (1984) “Stability and polarization of interests in job matching”, *Econometrica,* 52, 47-57.

Roth, A. and M. Sotomayor (1990) “Two-sided matching. A study in game theoretic modeling and analysis”, *Econometric Society Monographs No. 18, Cambridge University Press.*

Shapley, L. and M. Shubik (1972) “The assignment game I: The core”, *International Journal of Game Theory* 1, 111-130.

Sotomayor (1992) “The multiple partners game”, *Dynamics and Equilibrium; Essays in honor to D. Gale,* edit. by M. Majundar, published by Mcmillian.

______________(1999) “Three remarks on the many-to-many stable matching problem”, *Mathematical Social Sciences* 38: 55-70.

______________(2007) “Connecting the cooperative and competitive structures of the multiple-partners assignment game”, *Journal of Economic Theory* 134, 155 – 174

______________(2010) "Stability property of matchings is a natural solution concept in coalitional market games", *International Journal of Game Theory* 39:237–248.

______________(2012) "Labor time shared in the assignment game lending new insights to the theory of two-sided matching markets", *mimeo.*