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

by

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 describe 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 then proposed a new model, called deviation function form, which modifies Rosenthal's setting by also modeling the coalition structure and by incorporating new coalitional interactions, which support the agreements proposed by deviating coalitions. The concept of stability of the matching models, viewed as a cooperative equilibrium concept, is then translated to any game in the deviation function form and is confronted with the traditional notion of core. Precise answers are given to the questions raised.

Key words: cooperative equilibrium, solution concept, 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.

INTRODUCTION

The theory developed here presupposes an environment in which some decision situation takes place, involving a set of agents whose main activity is to form coalitions. If a coalition is formed, its members interact among them, by acting according to established rules, aiming to reach an agreement on the terms that will regulate their participation in the given coalition. This situation is called "cooperative game" when it is approached from an abstract mathematical view point. An *outcome* is the mathematical object which models a set of coalitions, whose union is the whole set of agents, together with the set of agreements reached by each coalition, as the final result of the negotiation process. Of course, agents have preference over possible outcomes[2].

A mathematical treatment for such a cooperative decision situation involves to simplify and to abstract the rules of the game. This is done through the built of an appropriate mathematical model. The starting point is the selection of the details of the rules of the game to be retained. Such selection depends on the purposes of analysis.

In this paper we propose a mathematical model for the situation described above, called *deviation function form* (*df*-form, for short). The rules of the game are fully modeled by the deviation function. For each feasible outcome x and coalition $S,$ the deviation function specifies the set of *feasible deviations from* x *via* $S.$ This set can be interpreted as the set of outcomes which can arise from the joint actions of the members of S against x, when these actions are allowed by the rules of the game.

This model can be used as vehicle of any cooperative equilibrium analysis, under the presupposition of the following principle of rationality:

"Faced with a feasible outcome x, a coalition will take any action against x, whenever such an action is allowed by the rules of the game and propitiates a profitable deviation from x to all members of the given coalition."

[2] An example of such a cooperative game situation is a matching market. The interested reader can find an introduction to the theory of matching markets in Roth and Sotomayor (1990, 1992).

Within this context, no cooperative equilibrium analysis should ignore "what coalitions can do *ex post"* when confronted with a feasible outcome, in order to deviate from it. According to this principle,

"*an equilibrium occurs when ex-post, after the agreements have been reached, there is no coalition, whose members, by acting according to the rules of the game, are able to profitably deviate from the set of current agreements."*

It is this principle that we shall try to formulate mathematically. Once this is done, a precise definition of the cooperative equilibrium, needed for rigorous reasoning, can be provided. Roughly speaking, a feasible outcome x is *destabilized* by a coalition S if there is some 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 a cooperative equilibrium if it is not destabilized by any coalition.

The framework presented here is general enough to model cooperative games in the normal form, in the characteristic function form or in the effectiveness form. It also encompasses situations in which a player may participate in more than one coalition. In these cases, when the trades done by a given coalition are independent of the trades done by the other coalitions, the *standard deviation*- to dissolve all current agreements and to form new agreements only among themselves – is not the only kind of deviation allowed by the rules of the game to the members of a coalition. The independent trades imply that the players of a coalition are allowed to maintain current agreements or to reformulate them. In such situations the core may not be the natural solution concept. (See Sotomayor 1992, 2010, 2012).

Historically, the cooperative equilibrium analysis of many games has relied on the notion of core and has had as its primary vehicle the characteristic function form. In this form an outcome is represented by the payoffs of the players, but the information with respect to the actions the players take to reach these payoffs is lost. For each coalition it is specified the set of payoff vectors that the members of the coalition can assure itself in some sense, without any concurs of players out of the coalition. Thus, the conditionality that characterizes the payoffs of the players in the cooperative games in the normal form is also lost.

In his paper of 1972, Rosenthal shows that in certain situations, in which the payoffs of the members of a coalition also depend on the actions taken by players out of the coalition, there may be outcomes (given by payoff vectors) in the core of the game in the characteristic function form, which are derived from outcomes (given by a profile of actions) that are not in the core of the cooperative game in the normal form. In an attempt to correct the imperfections of the characteristic function representation, which make this form inadequate to represent some cooperative games in the normal form, Rosenthal proposed the effectiveness function form.

Under the effectiveness function representation, for a given coalition and a given outcome there is a set of alternative outcome subsets, which the coalition can enforce against the given outcome. This means that when faced with an outcome, the members of a coalition are able, by interacting only among themselves, to restrict the negotiation process to any one of the specified subsets of the outcome set.

However, not all games fit equally well into the Rosenthal's framework, since the joint actions that a coalition can take *ex-post* are not always restricted to "interactions among themselves". Also, the cooperative equilibrium analysis based only on the core is not always the most appropriated approach. The point is that there may be some relevant features of some cooperative decision situations, which should not be ignored in any analysis of cooperative equilibrium of these situations, but that both game forms, the coalitional function form and the effectiveness form, seem to deal inappropriately.

In fact, in the text of the present work, an example presents a game situation whose representation in the coalitional function form or in the effectiveness form does not capture the details of the rules of the game that would be necessary to make evident if a given core outcome is or is not a cooperative equilibrium.[3] Thus the features of the game that are lost in these representations make deficient any cooperative equilibrium

[3] That there may be core outcomes that are not cooperative equilibria was first proved in Sotomayor (1992) (see also Sotomayor, 1999 and 2010).

analysis which has as vehicle any of the two forms. In this example, an adequate representation of the game situation is obtained by using the deviation function form. This model corrects the imperfections inherent of both, the coalitional function form and the effectiveness form, as those pointed out in this example.

It is well-known that in the matching models, the set of cooperative equilibrium outcomes (called stable outcomes in those models) is a subset of the core, and it may be smaller than this set.[4] We show that this set inclusion relation between the two sets is maintained in any cooperative game.[5] We define the effectiveness function form and the coalition function form for a game in the *df*-form. We then show that the effectiveness function form is adequate to represent cooperative games in which the coalition structure is a partition of the whole set of players. If, in addition, the payoffs of the members of a coalition only depend on the agreements reached by them, then these games can also be fully represented in the coalitional function form. In this case, when an outcome is represented by a payoff vector in both forms, the core of the game in the effectiveness function form coincides with the core of the game in the coalitional function form. In both cases, the concept of cooperative equilibrium is equivalent to the core concept.

Further details are discussed in the text. In section 2 the cooperative normal form, the coalitional function form and the effectiveness form are described and an example is presented with the intent of motivating section 3. In section 3 we present the deviation function form and give the definition of stability. Section 4 is devoted to model a cooperative decision situation in the *df*- form. Section 5 defines the core and the 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

[4] This phenomenom was first observed in Sotomayor (1992), for the continuous matching models and in Sotomayor (1999), for the discrete matching models.
[5] For the non-matching games, this phenomenom was first pointed out in Sotomayor (2010).

coalitional games,[6] the stability concept is equivalent to the core concept. Section 8 presents some final remarks.

2. PRELIMINARIES

The normal form of a cooperative game is derived from strategic situations in which agents can gain from the 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 $\sigma_j \in \Sigma_{Sj}$, $j=1,\ldots,k$, 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 are taken to 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]

For any vector $x \in R^n$, let x_S denote the projection of x on $R^{|S|}$. A domination relation on the set of feasible outcomes can be defined as follows.

> For x *and* y *in* R^n, *the vector* y ***dominates*** *the vector* x *via coalition* S *if (a)* $y_p > x_p \ \ \forall p \in S$ *and (b)* $y_S \in V(S)$. *Vector* x *is in the* **core** *of the game (N,V,H) if it is in* H

[6] Coalitional game is the name given by Shapley to a game that can be adequately represented by its characteristic function form.
[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 Aumann and Hart (1992), volumes I and II.

and it is not dominated by any vector in H via some coalition.

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 function $u: X \to R^{|N|}$; and (d) for each point $x \in X$, an *effectiveness function* E_x 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 coalition S is said to be *effective against* x, for any such subset of outcomes.

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 exist no objections.*

The idea of an objection is the following. Suppose an outcome x arises in a negotiation process. Suppose that coalition S, through actions that only involves players in S, enforces the set $Y \in E_x(S)$. 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

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

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 characteristic function form. Its main feature is that it is adequate to model cooperative games in normal form, since it captures game situations in which the utility levels reached 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, presents an outcome which is in the core of the game in the effectiveness form. However, there are coalitional interactions that can destabilize this outcome. This example also illustrates that there may be some features of the cooperative decision situations which the effectiveness form and the coalition function form seem to deal inadequately.

EXAMPLE 2.1. (The effectiveness form and the characteristic function form do not capture all relevant details of the rules of the game; an outcome is in the core of the game in the effectiveness form but it is not a cooperative equilibrium) 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 only 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 is not allowed to acquire more than 5 units. The market operates as follows. The agreements are negotiated by each seller and the buyer, individually. An agreement between one of the sellers and the buyer is independent of the agreement between the other seller and the buyer. A transaction between the buyer and any one of the sellers only occurs if the price of one unit of the good is any non-negative number less than or equal to \$3 and the number of units sold by the seller does not exceed the minimum between

[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.

the number of units that he owns and 5 minus the number of units negotiated between the buyer and the other seller. In this case each seller receives the product of his negotiation with the buyer. If buyer p acquires k items from a seller, for \$$t$ each, then he receives an individual payoff of $(3-t)k$ and the seller receives the payoff of tk.

Suppose that this market operates under a decentralized setting where the agents can freely communicate with each other. Any negotiation between the buyer and any of the sellers involves an agreement on the price and an agreement on the number of units to be negotiated. It is reasonable to assume that the whole agreement is broken once its terms with respect to the price are changed. However, as for the terms with respect to the number of items, two cooperative approaches take into consideration that different kinds of pairwise interactions may originate, depending on how the rules of the market open up for more flexible or more rigid agreements. Under rigid agreements, if the term with respect to the number of items is broken then the whole agreement is nullified[11]. A flexible agreement allows the buyer to decrease the number of units without breaking the agreement corresponding to the price.

Now consider the following feasible outcomes for the effectiveness form of the game. Under outcome x, buyer p acquires 5 units of the good of seller q_1 and pays \$2 for each unit. At the feasible outcome y, agent p acquires 4 units of the good of seller q_1, for the same \$2 for each unit, and q_2 sells his item to p for \$1.

Now observe that the point (5,10,0) 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 (6,8,1), which players p and q_2 both prefer. By using the characteristic function V one can only conclude that (6,8,1) is in $V(N)$ and (6,1) is not in $V(p,q_2)$, so (6,8,1) does not dominate (5,10,0) via coalition $\{p,q_2\}$. Indeed, the payoff-vector (5,10,0) is clearly undominated, so **it is in the core of the coalitional function form of the game.**

[11] This market is an instance of the time sharing assignment game introduced in Sotomayor (2011).

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**.

Now observe that the details of the market rules concerning the nature of the agreements are not modeled, either by the effectiveness form or by the characteristic function form. Therefore, there is no way to conclude from these representations, which kind of agreement is being used in the negotiations. Although the nature of the agreements can be neglected in an analysis based upon the core, it cannot be ignored if the focus is on the cooperative equilibria. which makes it inefficient any cooperative equilibrium analysis of this market based on any of the two forms. In fact, if the agreements are rigid, it is easy to verify that there is no way for p to increase his total payoff by only trading with q_2. In order to increase his total payoff, p must trade with both sellers, but there are no prices that can increase the current total payoffs of the three agents. Therefore, outcome x **is a cooperative equilibrium when the agreements are rigid.** Nevertheless, if the agreements are flexible and x is proposed, then buyer p and seller q_2 can counter-propose an alternative outcome that both prefer. 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. These actions are allowed by the rules of the market. The outcome y can be the resulting outcome if q_2 sells his item to p for \$1. The power of p of increasing his payoff is due to the concurs of q_1, 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 the agreements are flexible.**■

3. THE DEVIATION FUNCTION FORM AND THE SOLUTION CONCEPT OF STABILITY

In this section, motivated by Example 2.1, we propose the *deviation function form,* which provides a more general model than the effectiveness form for the purpose of cooperative equilibrium analysis. For a game represented in the deviation function form, we define a solution concept, which captures the intuitive idea of cooperative equilibrium. It will be called *stability*, due to the fact that restricted to the matching models that have been presented in the literature since Gale and Shapley (1962), it coincides with the concept of stability that has been established for them.

Basically, this representation is obtained when we abstract from the rules of the game and focus on the outcomes that result if players act according to these rules. Here are the primitives of this model.

- *(a)* *a* ***set*** $N = \{1, \ldots, n\}$ ***of players;***
- *(b)* *a* ***set*** C ***of coalitions;***
- *(c)* *For each set* $\mathcal{B}$ *of coalitions in* C *whose union is* N *(coalition structure), a set* $X_{\mathcal{B}}$ *of* ***feasible outcomes compatible with*** $\mathcal{B}$.
- *(d)* *for each* $p \in N$, *for each coalition structure* $\mathcal{B}$ *and* $S \in \mathcal{B}$, *with* $p \in S$, *a utility function* $U_{pS}: X_{\mathcal{B}} \to R$.
- *(e)* *An ordinal, vector-valued* **utility function** $u: X \equiv \cup X_{\mathcal{B}} \to R^n$.
- *(f)* *For each* $x \in X$, *a* **deviation function** ϕ_x, *which maps a coalition* $S \subseteq N$ *in to a set of feasible outcomes.*

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 set C can be interpreted as the set of **permissible coalitions**, that is, the set of non-empty subsets of N, whose elements can develop some joint activity allowed by the rules of the game. It is assumed that $N \in C$. We can think of the feasible outcomes in $X_{\mathcal{B}}$ as modeling the decisions (final agreements allowed by the rules of the game) that the players of each coalition $B \in \mathcal{B}$ can take with respect to their participation in B. 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.

Note that, unlike the normal form of a cooperative game, a coalition structure is not necessarily a partition of N and the set of agreements established by a coalition is not, necessarily, restricted to a set of actions or strategies. In the two-sided matching models, for example, a coalition structure is a feasible two-sided matching, which, in the discrete case, can also model the corresponding feasible outcome.

We can interpret $U_{pS}(x)$ as the utility level reached by p if he contributes to coalition S at the outcome x. The number $U_{pS}(x)$ is called *p's individual payoff at* x *via coalition* S. In many games, the utility level reached by a player via a coalition S does not depend on the agreements made by coalitions other than S.

Given an outcome x, the coalition structure compatible with x, together with the individual payoffs obtained at x by each player in each coalition he contributes, is called the *payoff configuration associated to* x. It is feasible if x is feasible.

The structure of preference for the players over the feasible outcomes is given by the utility function $u=(u_1,\ldots,u_n): X\to R^n$.

The functions ϕ_x are called **deviation function from** $\boldsymbol{x}$ and the feasible outcomes belonging to $\phi_x(S)$ are called **feasible deviations from** $\boldsymbol{x}$ **via** $\boldsymbol{S}$**.** These outcomes intend to reflect, in some sense, which feasible actions - that are allowed by the rules of the game - the members of S can take against x.

For each $y\in\phi_x(S)$ define the set $\phi^*_x(S,y)$ as follows. Let $\mathcal{B}$ be a coalition structure compatible with y. Set $S^*(y)=\{B\in\mathcal{B};\ B\cap S\neq\phi\}$. Then

> *$\phi^*_x(S,y)$ is the set of all feasible deviations from x via S that agree with y on $S^*(y)$.*

Clearly, $y\in\phi^*_x(S,y)$. This set can be interpreted as the set of outcomes that could result if S deviated from x "by doing the same sort of things it does at y". That is, when faced with x, the coalition S can take actions that could restrict the negotiation process to any one of the subsets $\phi^*_x(S,y)$'s. This does not mean that S can determine the particular outcome in $\phi^*_x(S,y)$ that will result.

The following example illustrates these concepts and the fact that the effectiveness function may be different from the deviation function.

Example 3.1. Consider an instance of the College Admission model where c_1 and c_2 are the colleges; $s_1,\ldots,s_5$ are the students; the quota of c_1 is 3 and the quota of c_2 is 2.[12] The only information about the preferences of the agents we will consider is that any partnership between a college and a student is acceptable to both agents. Let x be the feasible matching where c_1 is matched to $\{s_1,s_2,s_3\}$ and c_2 is matched to $\{s_4,s_5\}$. Let $S=\{c_1,s_4\}$. The set $\phi_x(S)$ is the set of feasible matchings that matches s_4 to c_1 or that leaves both agents unmatched. Therefore, matching y, where c_1 is matched to $\{s_1,s_2,s_4\}$, c_2 is matched to s_3 and has one unfilled position and s_5 is left unmatched, belongs to $\phi_x(S)$. The matching z, where c_1 is matched to $\{s_1,s_2,s_4\}$ and c_2 is matched to s_3 and s_5, also belongs to $\phi_x(S)$. The matching w at which c_1 is matched to s_4 and has two unfilled positions; c_2 is matched to s_1 and s_5; s_2 and s_3 are left unmatched, as well as the matching w' at which all players are unmatched, are also in $\phi_x(S)$. The set of partnerships in y whose intersection with S is non-empty is $S^*(y)=\{\{c_1,s_1\}, \{c_1,s_2\}, \{c_1,s_4\}\}$. The set $\phi^*_x(S,y)$ is the set of all feasible matchings at which c_1 is matched to $\{s_1,s_2,s_4\}$. Then, $z\in\phi^*_x(S,y)$ and $\phi^*_x(S,y)=\phi^*_x(S,z)$. Clearly, the matchings w and w' are not in $\phi^*_x(S,y)$.

Suppose that, faced with matching x, c_1 and s_4 take the following actions: c_1 leaves its partnership with s_3; s_4 leaves his partnership with c_2; c_1 and s_4 forms a partnership and c_1 retains its current partnerships with s_1 and s_2. In this case, the resulting outcome may be y, z, or any other matching in $\phi^*_x(S,y)$. However, coalition S cannot propose any outcome of $\phi^*_x(S,y)$ as an alternative outcome. Observe that the effectiveness function is different from the deviation function in this example: S is not effective for $\phi^*_x(S,y)$. In fact, S cannot *enforce* the set $\phi^*_x(S,y)$ against x, because the members of S are not interacting only among themselves at y. There are two subsets of $\phi_x(S)$ that can be

[12] See Roth and Sotomayor (1990) for an overview of this model.

enforced by S: $\phi^*_x(S,w)$ and $\phi^*_x(S,w')$. These two sets are the only ones for which S is effective. ■

In games in the deviation function form derived from cooperative decision situations, the intuitive meaning of a feasible deviation from an outcome x via some coalition S leads to the following natural assumptions.

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

(P2) *Let* $x'\in\phi_x(S)$. *If* $x\in X_{\mathcal{B}}$, $x'\in X_{\mathcal{B}'}$, $B'\in\mathcal{B}'$, $B'\cap S\neq\phi$ *and* $B'\not\subset S$, *then* $B'=B$, *for some* $B\in\mathcal{B}$ *and* $U_{pB}(x')\leq U_{pB}(x)$ *for all* $p\in B$.

(P3) *Let* x' *and* x'' *be feasible outcomes compatible with the coalition structures* $\mathcal{B}'$ *and* $\mathcal{B}''$ *respectively. Let* C' *and* C'' *be the sets of coalitions in* $\mathcal{B}'$ *and* $\mathcal{B}''$, *respectively, whose intersection with* S *is non-empty. If* $C'=C''$ *then* $x'\in\phi_x(S)$ *if and only if* $x''\in\phi_x(S)$.

(P4) $\phi_x(\phi)=\phi$

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

Assumption P1 says that if x' is a feasible deviation from x via S, which is compatible with the coalition structure $\mathcal{B}'$, then every player in S belongs to a coalition at $\mathcal{B}'$ formed only with players in S. P2 adds that 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 gets a utility level for his participation in B at x' that is higher than the utility level he gets for his participation in B at the current outcome x. P3 guarantees some *internal consistency.* P4 is a natural assumption. P5 implies that all outcomes at which the players in S interact only among themselves are feasible deviations from any outcome x via S.

The intuitive idea of cooperative equilibrium is that *an outcome x is a* **cooperative equilibrium** *if there is no coalition, whose members, by acting according to the rules of the game, can profitably deviate from x.*

The restriction of the concept of cooperative equilibrium to the matching models existent in the literature is captured by the concepts of stability already established for these models. **For this reason, this notion, translated to the deviation function form of a game, is captured by the solution concept defined below and called *stability*.**

For its definition we use a version of the domination relation introduced here.

Definition 3.1: *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) $y \in \phi_x(S)$.

Definition 3.2: *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.*

A payoff configuration is stable if it is associated to a stable outcome.

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

Let us consider an environment with a *finite set N of agents.* These agents are able to freely communicate and want to form *coalitions* (e.g. a matching market). In some situations, once a coalition is formed, the agents overtake an activity which generates some income which is split among them. In some of these cases, the division of the income among the agents is negotiated by them (e.g., the assignment game of Shapley and Shubik, 1972); in other cases, it is pre-fixed (e.g., the college-admission model of Gale and Shapley, 1962). More generally, the agents involved in a coalition try to make *agreements* (or to sign contracts) on the terms that will regulate their participation in that coalition. These terms may consist, for example, of the partnerships that should be formed, of the time of participation in the coalition, of the activity that should be developed, of the distribution of resources among the members of the coalition, of the actions that will be taken, of the choice of location, price or quality, etc.

There are *rules* specifying which actions a coalition can take and which ones it cannot take. Thus, for example, in a buyer-seller market the rules might require that a buyer is not allowed to acquire more than one object from the same seller and is not obliged to acquire any of the objects of a seller. In a labor market of firms and workers they might specify the total number of units of labor time available to each worker, and how a firm hires a given set of workers: in block or through individual and independent trades. If the trades are in block then any change in the group nullifies the whole trade; otherwise, a firm can keep some of the individual trades while dissolving some others, etc.

One can imagine that during the negotiation process, each agent tries to convince his potential partners that, in some sense, he is strong, by using his ability in showing that he has other, perhaps better, alternatives. Therefore, a sequence of "offers" and "counter offers" or "threats" and "counter threats", should culminate with a *feasible set of coalitions together with an agreement for each coalition*. The questions that naturally emerge are:

What coalitions should be formed? What agreements could be achieved?

It is reasonable to expect that the agreements that will be reached in a coalition should reflect in some sense the power of the members of that coalition. It then seems that a certain kind of *equilibrium* occurs if, *ex-post*, that is, after the agreements have been reached, *there is no set of actions that the members of a coalition "are allowed to take" to profitably deviate from the set of agreements obtained.*

A mathematical model for the cooperative decision situation which configures in such environment, if one abstracts from the negotiation process and focuses on the set of resulting outcomes, that is, if one focuses on what each coalition can obtain without specifying how to obtain, is what game theorists call *cooperative game.*

This session is devoted to the modeling of a cooperative game in the *df*- form, as defined in section 2. The *df*-form of a cooperative game can be used as vehicle for any equilibrium analysis based on the notion of stability.

In the representation of a cooperative decision situation in the *df*-form, the set of *players*[13] is the set N of participants. The collection of coalitions C is the set of non-empty subsets of N, which the rules of the game allow to be formed. The elements of C are called *permissible coalitions* (*coalitions*, for short).

A set of *feasible agreements* (agreements, for short) ∇_S is associated to each coalition $S \in C$. An agreement $\partial_S \in \nabla_S$ models a final coalitional interaction among the players belonging to S, whether and when S forms, and that is feasible to be reached if these players act according to the rules of the game. An agreement must specify the part of it that is due to each player in S. We will denote by ∇ the collection of ∇_S's, for $S \in C$.

A set $\mathcal{B}$ of permissible coalitions whose union is N is called *coalition structure*. 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* $\mathcal{B}$.

Once the final result x is reached in the negotiation process among the players, it may exist several pairs $(\partial;\mathcal{B})$ that can be used to represent x. The question is then

What pairs $(\partial;\mathcal{B})$ *should be used to adequately model an outcome?*

For the game-theorectic analysis purpose, it will be convenient to model an outcome as a pair $(\partial;\mathcal{B})$ such that all coalitions in $\mathcal{B}$ are *minimal* in the sense defined below.

[13] For simplicity of exposition we will refer to a player as "he".

Definition 4.1. Let $\mathcal{B}$ *be a coalition structure and let* ∂ be an *agreement structure for* $\mathcal{B}$*. Coalition* $S\in\mathcal{B}$ *is a* ***minimal coalition*** *at* $(\partial,\mathcal{B})$ *if its members cannot reach the part of* ∂_S *due to them by rearranging themselves in proper sub-coalitions (not necessarily pairwise disjoint) of* S*. If every coalition in* $\mathcal{B}$ *is minimal then* $\mathcal{B}$ *is called* **minimal coalition structure**.

Definition 4.2. *An* ***agreement configuration*** *is a pair* $(\partial;\mathcal{B}) \equiv ((\partial_1,B_1)\ldots, (\partial_k,B_k))$*, where* $\mathcal{B}$ *is a coalition structure and* ∂ *is an agreement structure for* $\mathcal{B}$*, such that* B_j *is a minimal coalition at* $(\partial,\mathcal{B})$*, for all* $j=1,\ldots,k$*.*

If $(\partial;\mathcal{B})=((\partial_1,B_1)\ldots, (\partial_k,B_k))$ is an agreement configuration, we say that $\mathcal{B}$ is *compatible* with ∂, and vice-versa. It must be clear that if $\mathcal{B}$ has the maximum number of coalitions among all coalition structures of pairs $(\partial;\mathcal{B})$'s, which can be used to represent a given outcome, then $\mathcal{B}$ is a minimal coalition structure, but the converse is not always true.

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.

In a matching market of firms and workers, for example, an agreement configuration is given by the matching, which specifies who works to whom, and the agreement structure, given by the corresponding arrays of payoffs, which specify the salaries a worker should receive from the firms which hire him.

Remark 4.1. It is worth to point out that the specification of the minimal coalitions in the description of an outcome allows, for instance, to distinguish whether a player interacts individually with each of his partners or in block. A consequence of Definition 3.1 is that, when agents make agreements in block, each agent enters only one minimal coalition.

If, say, 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 of the two-sided matching markets in which the players form multiple partnerships and the utilities are additively separable. In these markets, a minimal coalition at any outcome is formed by a pair of players from opposite sides or by single players. In the well known College Admission model of Gale and Shapley, a minimal coalition is formed by one student and one college, or only by a single student or a single college. ■

In the cooperative decision situations as those we want to model, it is reasonable to expect that an outcome, at which some coalition takes actions that are injurious to its own welfare, will never arise from a negotiation process among its members. E.g., one cannot expect to observe an outcome where a player gets less in some of the minimal coalitions he belongs than the minimum he can "assure" to himself by playing as a single player. The same way, one cannot expect to observe an outcome where a sub-coalition of a minimal coalition get less than it can get by playing in separate. More specifically, only outcomes at which the players take *rational decisions* and do not violate the rules of the game are feasible to occur. These are the *feasible outcomes*. Formally,

> *Under the assumption that the players take rational decisions and act according to the rules of the game, a* ***feasible agreement configuration*** *is a representation of a possible outcome of the game. The set of feasible outcomes compatible with 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.*

Given a feasible agreement configuration $(\partial;\mathcal{B})$ and a coalition $B \in \mathcal{B}$, the utility level enjoyed by an agent p belonging to B, at $(\partial;\mathcal{B})$, for his contribution to B, is given by a *utility function* $U_{pB}: X_{\mathcal{B}} \rightarrow R$. The number $U_{pB}((\partial;\mathcal{B}))$ is called *p's individual payoff at* $(\partial;\mathcal{B})$ *associated to* B.[14]

A minimal coalition structure, together with the individual payoffs obtained by each agent in each coalition he contributes is called *payoff configuration.* A payoff configuration is feasible if it is derived from a feasible outcome $(\partial;\mathcal{B})$. In this case we say that it is compatible with $\mathcal{B}$. In the Multiple partners assignment game of Sotomayor (1992), the partners must agree on the division of the worth of the pair and a player may contribute to more than one partnership. Then a player derives payoff in each partnership he enters. In this model, a feasible 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.

In the mathematical model we are building, the players have preferences over outcomes. A player compares two outcomes by comparing the corresponding payoff configurations; he compares two payoff configurations by comparing the corresponding arrays of individual payoffs; he compares two arrays of individual payoffs by comparing the level of utility he derives from each one. Thus,

the structure of preferences over the outcomes is given by an ordinal ***payoff function*** $\boldsymbol{u}$ *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 $u_p(x)=u_p(y)$.

[14] 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.

The translation of the definition of the set of **feasible deviations from an outcome** x **via coalition** S given in section 3 requires some preliminaries. Let S be a non-empty subset of N. A player is called a ***partner*** of S under an agreement configuration $y=(\partial,\mathcal{B})$, if he is member of a coalition $B\in\mathcal{B}$ such that $B\cap S\neq\phi$. If $S=\{p\}$, the partners of $\{p\}$ are all players belonging to the coalitions in $\mathcal{B}$ that contain p and so they may not be concentrated in the same minimal coalition.

We will use the notation P[S; y] to denote the set of all partners of S at y. Therefore,

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

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

Note that $S\subseteq$ P[S; y]; i.e., each member of S is also a partner of S.

Given a coalition S and an agreement configuration $y=(\partial,\mathcal{B})$, the sets $S^*(y)$ and $\phi^*_x(S,y)$ are translated to our model as follows.

$$S^*(y)\equiv\{B_j\in\mathcal{B};\ B_j\subseteq P[S;y]\} \text{ and } \partial_{S^*(y)}\equiv\{\partial_{Bj}\in\partial;\ B_j\in S^*(y)\}.$$

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

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

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 we can decompose y in two parts and represent it as $y=(y^{S^*}, y^{\mathcal{B}\backslash S^*})$.

In the cooperative decision situations we want to model it is assumed the following principle:

Confronted with a feasible outcome x, the players of a coalition S will "deviate from x" by taking actions against x (even in case these actions involves current partners out of the coalition), allowed by the *rules of the game, whenever any one of the resulting outcomes can benefit all players in S. On the other hand, no coalition of players is able to profitably deviate from a cooperative equilibrium outcome, by acting according to the rules of the game.*

The several types of coalitional interactions the members of S are allowed to perform against x, in order to deviate from x, are specified by the rules of the game situation. They lead to the set $\phi_x(S)$ of *feasible deviations from* x *via* S. If $x \in X$ and $y \in \phi_x(S)$, the set $\phi^*_x(S,y)$ is the set of *feasible deviations from* x *via* S *generated by* y^{S^*}. That is,

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

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

Faced with x, if coalition S deviates from x according to y^{S^*}, $\phi^*_x(S,y)$ is the set of outcomes that 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. We say that S *destabilizes* x if all members of S prefer *every* feasible deviation of $\phi^*_x(S,y)$ to x. The feasible outcome x is *stable* if it is not destabilized by any coalition. Therefore, the cooperative equilibria for the cooperative games are the *stable outcomes* defined in section 3.

Thus, in order to check instabilities we have to know the kinds of deviations which are feasible (and may be expected) to a given coalition 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)-(P5) 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 $x\in X$ and $S\in C$, for all $y\in\phi_x(S)$, there is no minimal coalition 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 $y\in\phi_x(S)$, the players of S only perform *standard coalitional interactions against* x: they discard all current agreements at x and perform some $\partial_S\in\nabla_S$, compatible with a new 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 S deviates from x, the agreements chosen by the players in $N\backslash S$ do not affect the new utility levels of the players in S. Thus, the feasible deviations of S from x are independent of x and are given by:

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

Under this approach the coalitional function is given by:

$$V(S)=\{u\in R^{|S|};\; u_p=U_{pB}(y),\text{ with } y\in\phi(S),\, B\in S^*(y) \text{ and } p\in B\}$$

The other approach is appropriated 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 $\partial_j=\gamma_k$.

Example 4.1, at the end of this section, illustrates 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.

2nd case: *The coalition structures associated to the outcomes are not, necessarily, partitions of* N *and the agreements 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, the agreement between the buyer and a seller only concerns the price of one unit of the good. Once this price is set, the number of units demanded by the buyer is always accepted if this demand can be satisfied by the seller. It might then be the case that the buyer wanted to reduce the number of items that was being proposed at x, by keeping the current price of one item. Coalitional interactions of this sort are called *agreement reformulations*. We can model this type of flexibility in Example 2 by considering that the term on the price and the term on the number of items to be negotiated are independent.

In general, when agreements are flexible, some players may want to keep some of its current partnerships, which contain partners out of S, and to reformulate some of the current agreements of these 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 not only an agreement different from γ_j, but it is also an agreement that is not a reformulation of γ_j, so at least one player in D_j prefers ∂_j to γ_j.

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 this case, $\phi^*_x(S,y)$ *is the set of feasible outcomes* y' *which agree with* y *on* $S^*(y)$.

Note that in (i) it is not required that $B_j \not\subset S$. In Example 2, outcome y is a feasible deviation from x via $S=\{p,q_2\}$ when the agreements are flexible. There, as well as in Example 1, the power of coalition S also depends on interactions among some of its members with some current partners out of the coalition. ■

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

Under a rigid agreement, if some of the terms is 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. Then it is possible that the members of S arrange themselves into a feasible set of minimal active coalitions (i) by discarding some current minimal active 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* $B_k \subseteq S$.
In this case, $\phi^*_x(S,y)$ *is the set of feasible outcomes* y' *which agree with* y *on* $S^*(y)$.

Note that in (i) it is not required that $B_j \not\subset S$. ■

When the agreement structures of the feasible outcomes are given by arrays of individual payoffs, one array for each player, then if $y=(u,\beta)$ is a feasible deviation from $x=(v,\gamma)$ via S, we say that u is a feasible deviation from v via S. In this case we also say that u *is stable* if the corresponding agreement configuration (u,β) is stable.

The following example illustrates these definitions.

EXAMPLE 4.1. Consider a game in the deviation function 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 $U(y=(\partial;\beta)$ 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 $\phi_x(S)=\{y,z\}=\phi^*_x(S,y)=\phi^*_x(S,z)$, where $y=(\{\sigma_1\},\{\gamma_2\},\{1\},\{2\})$ and $z=(\{\sigma_2\},\{\gamma_2\},\{1\},\{2\})$. We have that $U_2(x)=3$, $U_2(y)=4$ and $U_2(z)=2$. Then, $U_2(y)>U_2(x)>U_2(z)$, so x is ϕ-dominated by y via S and is not ϕ-dominated by z via S. Hence, $\{2\}$ does not destabilize x according to Definition 2.2. 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 the players are faced with this game, they 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 $\phi_x(S)=\phi^*_x(S,y)=\{y\}$. Under this approach, $\{2\}$ destabilizes x and so x is not stable. ■

In a wide class of games, which includes many of the matching models, given $y \in \phi_x(S)$, the payoffs of the players in S only depends on the part y^{S^*} of y, so the players in S are indifferent between any two outcomes in $\phi^*_x(S,y)$. This class of games will be denoted by G^*. 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 $\phi^*_x(S,y)$. Then we can rewrite Definition 2.2 as follows:

> **Definition 4.3:** *An outcome* x *is* ***stable*** *for the game* $(N,X,U,u,\phi) \in G^*$ *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 belonging to* G^* *is equivalent to the definition of stability that has been established for these models.*

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

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

$E\phi_x(S) \equiv \{y=(\partial,\mathcal{B}) \in X;\ y \in \phi_x(S)$ and if $B \in \mathcal{B}$ then either $B \subseteq S$ or $B_j \subseteq N \backslash S\}$
(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, so $P[y,S]=S$. 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 in $E\phi_x(S)$ may also depend on the agreements made by players in $N \backslash S$. Property P5 asserts that

$E\phi_x(S) \subseteq \phi_x(S)$ for all feasible outcome x. In the situation described in Example 3.1, $E\phi_x(S)= \phi^*_x(S,w)\cup\phi^*_x(S,w')$.

Given $\Gamma=(N,C,X,U,u,\phi)$, 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 must be understood that the effectiveness function form of a game fully represents this game if and only if $E\phi_x(S)=\phi_x(S)$ for every $x\in X$ and every $S\subseteq N$.

If $x\in X$ and $y\in E\phi_x(S)$, the set of feasible deviations from x via S according to y^{S^*} can be identified with the set of outcomes that is enforced by S against x via y, which will be denoted by $E^*\phi_x(S,y)$. That is,

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

The feasible outcome x is in the **core** of a game in the *df*-form if there is no coalition S and no $y\in E\phi_x(S)$, such that any outcome that can be enforced by S via y is preferred by all players in S to the current outcome x. Formally,

Definition 5.1: *Let $\Gamma=(N,C,X,U,u,\phi)$ be a game in the deviating 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 5.2: *Let $\Gamma=(N,C,X,U,u,\phi)$. 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 (N,C,X,U,u,ϕ) if it is not blocked by any coalition.*

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

Definition 5.3: *Let $\Gamma=(N,C,X,U,u,\phi) \in G^*$. An outcome is in the* ***core*** *of Γ if it is feasible and it is not dominated by any feasible outcome via some coalition.*

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

Theorem 6.1 provides the link between the core and the stability concepts in a game in the *df*-form: the stability concept can be viewed as a refinement of the core concept. Theorem 6.2 assures the equivalence between the two concepts when the effectiveness function form that can be derived from the *df*-form of the game can adequately represent this game.

Theorem 6.1. *In a game $\Gamma=(N,X,\phi,u)$, the set of core outcomes contains the set of stable outcomes.*

Proof. It is immediate from Definitions 5.1, 5.2, 2.1 and 2.2 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. *Suppose the coalition structures associated to the feasible outcomes for a game $\Gamma=(N,C,X,U,u,\phi)$ are partitions of N. Then, $E\phi_x(S) = \phi_x(S)$, for any feasible outcome x for Γ.*

Proof. It is implied by (P5) that $E\phi_x(S)\subseteq\phi_x(S)$ for any feasible outcome x. Thus we only have to show the inclusion in the other direction. It follows from the hypothesis that the players are allowed to enter one minimal coalition at most. Therefore, (P1) implies that if $y\in\phi_x(S)$ then there is no minimal coalition at y that contains elements of S and elements of $N\backslash S$. It is then implied by (P5) 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* $\Gamma=(N,C,X,U,u,\phi)$. *Suppose that, for every coalition* S *and feasible outcome* x *for the game* Γ, $E\phi_x(S) =\phi_x(S)$. *Then, in* Γ, *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 2.2 implies that there is 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 5.1 that x is dominated via S by every outcome in $E\phi^*_x(S,y)$, so Definition 5.2 implies that x is blocked by S, which is a contradiction. The other direction follows from Theorem 1. Hence the proof is complete. ■

Corollary 6.1. *Suppose the coalition structures associated to the feasible outcomes for a game* $\Gamma=(N,C,X,U,u,\phi)$ *are partitions of* N. *Then, in this game, 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 6.1 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 active coalition. For this model, it is easy to construct examples in which $E\phi_x(S)\neq\phi_x(S)$ for some outcome x.

7. COALITIONAL GAMES

The deviating 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 $u_S \in V(S)$, it can be interpreted that S can take some joint action which yields itself at least u_S. This payoff 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 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 minimal coalitions. Thus a feasible outcome is a payoff configuration (u,∂) such that $u \in V(N) \subseteq R^{|N|}$ and ∂ is a partition of N.

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, let X be the set of the feasible outcomes. Define

$E\phi(S)=\{(u,\partial)\in X;\ u_S \in V(S)\}$

From our intuitive discussion of feasible deviations via a coalition S we should require that $E\phi(S)\equiv\phi_x(S)$, for every feasible outcome x.

Since every coalition structure is a partition of N, Corollary 6.1 implies that the core equals the set of stable payoffs in $\Gamma=(N,C,X,U,u,\phi)$ and so does in (N,V). Therefore we have proved that

Theorem 7.1. *Let (N,V) be a coalitional game. Then, in this game, the set of core payoffs equals the set of stable payoffs.*

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

Let x and y be feasible outcomes for the *df*-game $\Gamma=(N,C,X,U,u,\phi)$ associated to the coalitional game (N,V). Clearly, 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. The stability definition can then be rewritten as follows:

> **Definition 7.1:** *A feasible payoff-vector u 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 $v_i>u_i$ for all $i\in S$ and $v_S\in V(S)$.*

This is the usual concept of core for coalitional games.

8. FINAL REMARKS

Following the approach of Gale and Shapley, some attempt has been done in the mathematical modelling of the matching markets presented in the literature, in the sense that the concept of stability be established as the concept that 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 matching markets are mathematically modeled as cooperative games.[15]

However, the concept of stable outcome has been locally defined for each matching model that has been studied. It should not then be surprising that in this process, the definition of stability has not always

[15] 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 as a cooperative game in the coalitional function form.

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 decides that the outcomes with certain properties will be called "stable". The intuition behind the definition is not discussed. Therefore, a general definition of stability, which establishes the concept of stability for every matching model, fills an important gap in the literature.

This work grew out of the attempt to formulate mathematically the following principle, on which we believe the theory of cooperative equilibrium should be supported: *"faced with a feasible outcome x, a coalition will take any action against x (even in case this action involves current partners out of the coalition), whenever such action is allowed by the rules of the game and propitiates to all members of the given coalition a more profitable outcome than x; on the other hand, no coalition of players is able to profitably deviate from a cooperative equilibrium outcome, by acting according to the rules of the game".*

The mathematical formulation of this principle led to the concept of cooperative equilibrium. In order to properly define it we observed that the intuitive idea of stability for matching models would also apply to a non-matching game if, instead of pairwise interactions the players interacted with coalitions of any size. We then brought out a new form of representing those games whose feasible outcomes can be supported by a coalition structure of minimal coalitions. In this model, the kinds of actions that a coalition is allowed to take against an outcome are naturally expressed by a deviation function. We showed that some features of certain cooperative game situations, which can be expressed under the deviation function form, may fail to be captured in the coalitional function form and in the effectiveness form, yielding incorrect conclusions in both game forms. We analyzed the correlation between the set of core outcomes and the set of stable outcomes. The main result is that, in any cooperative game, the stability concept is a refinement of the core concept and it is stronger than the core concept, as it happens in the matching models.

In practical terms, the theory developed here provides a mathematical model to represent adequately a variety of cooperative decision situations. From the conceptual point of view, to focusing on the stability concept rather than on the core concept, propitiates that new subjects of investigation be certainly created.

The idea of making accessible the matching theory to games which are endowed with a coalition structure, not necessarily a matching structure, was first considered in Sotomayor (2010-b). The concept of stability was defined by the first time for a class of games, where utilities are additively separable and, unlike the matching games, the coalition structure is given by coalitions of any size. In these games the set of stable outcomes may be a proper subset of the set of core 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.

Aumann, R. and M. Maschler (1964) “The bargaining set for cooperative games”, *in Annals of Math. Study 52* (Dresher, Shapley and Tucker, Eds.). 443-476, Princeton Univ. Press, Princeton, NJ.

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

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.*