Fifty Years of the Nash Program, 1953-2003*

Roberto Serrano†

Department of Economics, Box B
Brown University
Providence, RI 02912, U.S.A.

Working Paper No. 2004-21
December 2004

Abstract. This paper is a survey of the work in the Nash program for coalitional games, a research agenda proposed by Nash (1953) to bridge the gap between the non-cooperative and cooperative approaches to game theory.
Journal of Economic Literature Classification: C71, C72, C78.
Keywords: Nash program, cooperative games, non-cooperative games, bargaining, implementation.

*This paper is an extended version of the VI "Investigaciones Económicas" Lecture, delivered at the XXVIII Spanish Symposium on Economic Analysis, in Sevilla on December 12, 2003. I acknowledge support from NSF grant SES-0133113.

†Brown University, phone (401)863-2764, fax (401)863-1970, `roberto_serrano@brown.edu; www.econ.brown.edu/faculty/serrano`

1 Introduction

Fifty years have passed since the publication of Nash's article "Two Person Cooperative Games" (Nash (1953)). This important paper set a new entire research agenda that has been referred to as the Nash program for cooperative games. Much work in the Nash program has taken place in this half a century, and with the aid of this time perspective one should make an effort to summarize what we have learned. In writing this survey, I may be forgetting inadvertently some valuable contributions, and I do apologize in advance to their authors.

Similar to the microfoundations of macroeconomics, which aim to bring closer the two branches of economic theory, the Nash program is an attempt to bridge the gap between the two counterparts of game theory (cooperative and non-cooperative). This is accomplished by investigating non-cooperative procedures that yield cooperative solutions as their equilibrium outcomes. Nash himself stated his goal eloquently:

> "We give two independent derivations of our solution of the two-person cooperative game. In the first, the cooperative game is reduced to a non-cooperative game. To do this, one makes the players' steps in negotiations in the cooperative game become moves in the non-cooperative model. Of course, one cannot represent all possible bargaining devices as moves in the non-cooperative game. The negotiation process must be formalized and restricted, but in such a way that each participant is still able to utilize all the essential strength of his position. The second approach is by the axiomatic method. One states as axioms several properties that would seem natural for the solution to have, and then one discovers that the axioms actually determine the solution uniquely. The two approaches to the problem, via the negotiation model or via the axioms, are complementary. Each helps to justify and clarify the other."
> (Nash (1953, p. 128)).

Supporting Nash's words, Harsanyi (1974) elaborates on the goals of the Nash program and writes: "Nash (1953) has suggested that we can obtain a clear understanding of the alternative solution concepts proposed for cooperative games and can better identify and evaluate the assumptions to make about the players' bargaining behavior if we reconstruct them as equilibrium points in suitably defined bargaining games, treating the latter formally as non-cooperative games." The idea is both simple and important: the relevance of a concept (a cooperative solution in this case) is enhanced if one arrives at it from very different points of view. The purpose of science is to uncover "relationships" between seemingly unrelated concepts or approaches (see Aumann (1997)). To cite Nash (1953) again: "It is rather significant that this different approach yields the same solution. This indicates that the solution is appropriate for a wider variety of situations than those which satisfy the assumptions we made in the approaches via the model."

Indeed, the outcome of a successful result in the Nash program will be an enhanced understanding of a cooperative solution. Apart from the meaning attached to its definition and, in many cases, the normative content provided by the axioms that characterize it, one will have now a better idea of the institutional arrangements and negotiation protocols that lead to the solution as a consequence of self-interested -non-cooperative- rational play. As we shall see, it is very instructive to note the differences in the procedures that connect with the different cooperative solutions. Although the theory has not investigated yet necessary conditions on a procedure to arrive at a given solution, this is a development that may take place in the years to come. For now, what one can do is to underline some of the salient aspects of the mechanisms that have been used, trying to extract from them some properties that may be inherent to the solution. Before discussing the Nash program itself, I shall begin with a brief section on cooperative game theory.

2 Cooperative or Coalitional Game Theory

A number of authors have provided different definitions of cooperative game theory. Each of them has its merits, although for most of them we shall argue that they do not necessarily correspond to an exact definition; this has sometimes led to misunderstandings. We shall review them briefly here.

- Cooperative game theory has been sometimes portrayed as a theory about fairness. While this is certainly true for some cooperative solutions, such as the Shapley value or the Nash solution, because of the symmetry axiom behind them, it does not provide a good definition of the entire field. For example, the core, which is one of the leading solution concepts, has nothing to do with fairness, as it describes an environment of potential conflict among coalitions. It is not hard to see how there are core payoffs in many games that one would have difficulties in describing as "fair" distributions of the surplus.

- Some other times cooperative game theory has been defined as a theory based on an outside enforcement authority. According to this view, the agreements in the non-cooperative theory are self-enforcing (as in a Nash equilibrium), whereas payoffs in a cooperative solution are sustained by a binding agreement guaranteed by an outside enforcer. Two comments are in order here. First, the enforcement of such payoffs relates to their implementation and hence falls outside of the realm of cooperative theory: the theory identifies or recommends certain payoffs following some desirable normative principles, but is silent about how to impose them. And second, there is a sense in which non-cooperative theory also relies on outside enforcement. Namely, when the two prisoners play "Defect" in the prisoners' dilemma, a specific outcome takes place and is enforced. In some sense, I believe that the need of the enforcer is there for both counterparts of game theory (the entire neoclassical paradigm in economic theory presumes the existence of a well defined legal system that protects property rights and upholds and enforces contracts). To the extent that the Nash program relates to implementation theory, enforcement issues are always part of the approach. We shall come back to this in the sequel.

- A third definition is that of a normative theory based on axioms. It is true that many cooperative solutions have been characterized axiomatically. However, the axiomatic method in economics and game theory is not confined to this field. For example, the derivation of different non-cooperative equilibrium concepts (Nash equilibrium, correlated equilibrium) as consequences of different assumptions on rationality and knowledge are beautiful examples of applications of axiomatic methodology. Other important fields, such as individual decision theory and the theory of social choice, also employ the axiomatic method to conduct their inquiries.

- Is cooperative game theory that part of game theory that studies cooperation? The name of the field may be misleading in this respect. First, the cooperative approach in game theory need not assume that it is always the grand coalition the one that is formed; in general, the cooperative solutions can be adapted to accomodate arbitrary coalition structures, in which different coalitions compete with one another. In addition, cooperation is also a fundamental topic in the non-cooperative theory through the important insights provided by the model of repeated games. It is perhaps for these reasons that some authors (e.g., Aumann (1987)) advocate for the name *coalitional* instead of *cooperative* game theory.

- In my opinion, perhaps the most complete definition of cooperative game theory is that of a theory in which coalitions and the set of payoffs "feasible" for each coalition are the primitives. When given a strategic situation, one can begin by describing it in the most informative way possible, which corresponds to the *extensive form* representation. This is a game tree, specifying in great detail the timing and set of moves available and the information held by each player at each point in time. One step of abstraction is gained by reducing the extensive form to its *normal form* or *strategic form* representation: in it, we endow each player with a set of strategies, with independence of whether the strategy is to be executed following a specific timing of moves and under some informational constraints.

  Yet another step of abstraction is taken if one suppresses information concerning strategies, and is given simply the set of payoffs that each player (and each subset of players) can achieve as a function of the coordination reached by the rest of the players. In the most general form, the *partition function form* of a game expresses the set of feasible payoffs to a coalition for each coalition structure -partition- of the complement. Often a simplifying assumption is made, by which the set of feasible payoffs to a coalition is independent of the organization of the complement, and this leads to the *characteristic function form* representation. Most of the contributions in the Nash program have worked with the latter representation, although there are interesting exceptions that have employed the model of the partition function (e.g., Bloch (1996)). Thus, according to this definition, a coalition is a "black box." In identifying it with a set of feasible payoffs, one lacks a description of the sort of arrangements or procedures that the coalition is supposed to follow to achieve them. Nonetheless, the approach has a clear advantage in terms of robustness: indeed, the attempt is to derive recommendations -solution concepts- that are independent of the

> unimportant details of different procedures that underlie the same set of feasible payoffs. Note also how in this formulation we are silent about the behavioral assumptions that players obey. In particular, while fairness considerations will be sometimes used, following the imposition of the symmetry axiom and alike, the approach is also compatible with fully strategic players that embark in the formation of coalitions.
>
> Thus, we shall talk about non-cooperative game theory as that form of analysis that uses both the extensive and the strategic form representations of games, whereas cooperative game theory utilizes the partition and characteristic function forms, in which many details of the interaction have not been described explicitly.

Accepting the latter view, the role of the Nash program can be understood as opening the "black box" of each coalition. This is done by breaking down a solution payoff (or set thereof) based on a coalitionally-based definition into individual decision problems, built in suitably defined negotiation games.

We shall close this section with some definitions that will be used in the rest of the paper.

Let $N = \{1, \ldots, n\}$ be a finite set of players. Let $G = \langle (M_i, u_i)_{i \in N}$ be an n-player game, where M_i is the set of pure strategies of player i, and $u_i : \prod_{i \in N} M_i \mapsto \mathbb{R}$ is player i's payoff function.[1] If the game is one in extensive form, we would be given the game tre specifying the timing of moves and information sets for each player, from which one can derive the strategy sets.

Given a set of outcomes A, a mechanism or game form is defined as a pair $\langle (M_i)_{i \in N}, g \rangle$, where M_i is agent i's strategy set and $g : \prod_{i \in N} M_i \mapsto A$ is an outcome function. A mechanism can be designed as an institution to govern the interactions among agents, even if one does not know their preferences over outcomes. When the agents in N interact in the mechanism, each agent i will evaluate outcomes according to her payoff function $u_i : A \mapsto \mathbb{R}$. Therefore, composing a mechanism with a profile of payoff functions $(u_i)_{i \in N}$ one obtains a game.

A coalition S is a non-empty subset of N. For each coalition S, let $V(S) \subset \mathbb{R}^S$ be the set of payoffs feasible for S. We denote by $\mathbb{R}^S$ the Euclidean space with as many dimensions as players in S, which are labeled with the names

[1] For the most part, mixed strategies have been excluded from the analysis. We shall also do it here.

of the players in S. The sets $V(S)$ are assumed to be non-empty and closed. At times one also assumes boundedness; or in alternative specifications that allow free disposal of utility, these sets are assumed to be comprehensive, i.e., if $y \in V(S)$, and y' is such that $y'_i \leq y_i$ for every $i \in S$, then $y' \in V(S)$. The pair (N, V), where V assigns a set $V(S)$ to each coalition S is called a game in characteristic function. In this paper, this is what we shall take as a cooperative or coalitional game. Let Γ be a class of such games.

A cooperative solution F assigns to each game (N, V) in the class Γ a set of payoff profiles $F(N, V) \subset \mathbb{R}^N$ such that for each $x \in F(N, V)$, there exists a coalition structure (partition of N) $(S_1, \ldots, S_k)$, satisfying that for each $j = 1, \ldots, k$, $x^{S_j} \in V(S_j)$. Most of the work in the Nash program has assumed the formation of the grand coalition. Thus, in this paper, we shall simply work with solutions such that for each $x \in F(N, V)$, $x \in V(N)$. We shall use f to denote single-valued solutions.

3 The Nash Program and Implementation Theory

In viewing the Nash program as an agenda within the theory of implementation, one should at least describe briefly the main elements of this theory.[2] Implementation theory is concerned with the design of mechanisms to decentralize the decision-making process in any social choice problem.

A social choice problem consists of four components:

- (1) a set of agents,
- (2) a set of socially feasible alternatives or outcomes,
- (3) agents' preferences over those outcomes, and
- (4) a social choice correspondence, specifying a set of desirable outcomes as a function of the agents' preferences.

Examples of social choice problems include exchange and production economies, voting problems, and also cooperative games. To see that a

[2]See Moore (1992), Corchón (1996), Jackson (2001), Maskin and Sjostrom (2002), Palfrey (2002) and Serrano (2004) for surveys of the theory of implementation; see Hurwicz (1994) for a brief expository piece.

cooperative game fits naturally into this definition, one can list down its components:

- (1) a set of players N,
- (2-3) a characteristic function $V = (V(S))_{S \subseteq N}$, specifying a set of feasible payoff profiles for each coalition, and
- (4) a solution F, assigning a set $F(N, V)$ of prescribed payoffs to each cooperative game (N, V).

It is apparent that the only departures from the definition of a social choice problem are two: first, the specification of possibilities for each group of agents, not only for the whole set; and second, the fact that the "black box" of the characteristic function compresses in it considerations both over feasible outcomes and preferences over them. Indeed, underlying the set $V(S)$ one may have different sets of feasible outcomes, such as the redistributions of the initial endowment of goods among agents in the coalition, a production set describing the technological possibilities of the group, and so on.

Given a social choice problem, the goal of implementation theory is to design for the agents (element (1)) a mechanism using the feasible outcomes (2) so that, regardless of the agents' preferences (3), one ends up implementing the desirable outcomes (4) as a consequence of the agents' rational actions in the mechanism. What makes the exercise non-trivial is that the mechanism designer not know the agents' preferences and therefore, the mechanism itself cannot depend on that information. Yet the mechanism must "work" over a reasonably large class of social choice problems to be deemed a success.

Similarly, in the Nash program a mechanism is to be played by the set of players (1). However, the modeller has now an additional degree of freedom, since the characteristic function does not imply a fixed structure of feasible outcomes; one can write down different mechanisms based on different sets of physical or abstract outcomes (see Serrano (1997a), Dagan and Serrano (1998), Bergin and Duggan (1999), and Trockel (2002a, b) for different ways to endow the characteristic function with an outcome structure). Fixing the structure of outcomes, one would like to design a mechanism that, whatever the players' payoffs (i.e., the characteristic function), is capable of yielding the outcomes that, when evaluated by the players' payoff functions, lie in the solution. Again, one way to measure the success of an implementation result in the Nash program is to assess the class of cooperative games over which the result is obtained.

It is true that many contributions in the Nash program write down a mechanism whose rules depend on the characteristic function, but those mechanisms become more useful if their rules can be adapted in terms of outcomes, thereby making them independent of preferences. In particular, this means that if the characteristic function is known to the designer and the only goal of the Nash program were seen as suggesting a non-cooperative game whose set of equilibrium payoffs is $F(N, V)$, one can achieve this by the trivial mechanism: if all players agree on a given payoff profile in the solution, implement it; otherwise, implement a bad outcome. This point, which was in the folklore of the practicioners of the Nash program for a while, was fleshed out by Bergin and Duggan (1999, Proposition 1). However, this mechanism is not useful and does not shed any light on the meaning of the specific solution (as such, the mechanism "works" for any solution).

A similar objection can be raised in the case of solutions such as the egalitarian bargaining solution. As explained in Dagan and Serrano (1998), because this solution fails to be invariant with respect to equivalent representations of utilities over lotteries, any mechanism used for its implementation (Bossert and Tan (1995)) must be necessarily changing when one goes outside the class of bargaining problems over which the physical outcome associated with the solution is unchanged (e.g., a given bargaining problem and all those derived from it by multiplying the utility functions of each player by the same constant). In general, it is desirable that the mechanisms to be used not depend on agents' preferences; this makes them far more useful. This is one of the main advantages of liberating the mechanisms of the Nash program from their dependence on the characteristic function.

It has sometimes been argued that the Nash program and implementation theory are different agendas because, while the presence of a planner is necessary in the latter, it is not in the former. I believe this is not a substantial difference. First, it is not true that implementation theory requires the existence of a planner: the community of n agents involved in the problem at hand can be the one in charge of the design task. In addition, in undertaking the Nash program's endeavor, not only must one specify the details of the negotiations as moves in a mechanism, but one must also make the assumption that a given profile of moves leads to an outcome that can be some how enforced. Again citing Nash (1953): "The point of this discussion is that we must assume there is an adequate mechanism for forcing the players to stick to their threats and demands once made, for one to enforce the bargain once agreed. Thus, we need a sort of umpire who will enforce contracts or commitments." That is, the vexed issue of enforcement of outcomes cannot be overlooked, and one must assume either an enforcement by the designer or by other vehicles which will typically be left unspecified. Therefore, both in terms of who designs mechanisms and who enforces their outcomes, the Nash program does not depart from the assumptions made in the general theory.

The main question that the general theory of implementation has been concerned with is the identification of conditions on social choice rules and domains that allow implementability. Because of this general scope, the mechanisms proposed there are necessarily abstract. This is bound to change when the theory turns more to applications. One good example is precisely the mechanisms in the Nash program. A mechanism in the Nash program, by targetting a specific cooperative solution whose implementation is sought, should provide useful information about the types of negotiation protocols that lead to that solution. The consequence is an increased understanding of the solution in question, beyond the content of its definition and the axioms behind it.

On the other hand, one should not overstate the case. A typical result in the Nash program –say, the finding that a specific bargaining procedure yields the Nash solution as the unique equilibrium outcome– is sometimes described as a non-cooperative "foundation" of the cooperative solution; and some authors have pushed this view further, by stating that the relevance of a cooperative solution depends mainly upon such a finding (e.g., Muthoo (1999, Chapter 2)). My use of the quotation marks around the word "foundation" in the earlier sentence is deliberate. One tends to associate with that word a sense of robustness and imperturbability that is often lacking in an extensive form. That is why I am usually reluctant to the use of the phrase "non-cooperative foundations." On the other hand, for most cooperative solutions, an axiomatic system has been proposed that characterizes them. In my view, the normative principles embodied in the system of axioms behind a solution fits better the notion of a "foundation." That is not to say that the exercise of the Nash program is unimportant: one can learn much from the uncovery of the negotiation procedures that lead to a solution, i.e., from its non-cooperative implementation.

4 Pure Bargaining Problems

In this section we begin to survey specific results in the Nash program, starting with those obtained for the class of pure bargaining problems. A pure bargaining problem is a cooperative game in which only the grand coalition N can create a positive surplus with respect to what each player can achieve if she does not cooperate with anyone. That is, subcoalitions have no power; the agreement of all parties involved is necessary to the creation of a "pie." If the problem involves only two players, these are bilateral monopoly problems, in older jargon (i.e., a firm and a union bargaining over wages, a buyer and a seller negotiating over price). If there are more than two players, examples of pure bargaining problems include cooperatives of workers, all of whom are part of the firm, or budget negotiations among a group of countries, where each country may have veto power over agreements.

Let us recall the relevant notation, which we already introduced:

- $V(N)$: utility possibility set for the players in N. Typically, one assumes a lower bound to utility through normalization, so that $V(N)$ is compact. In addition, if the underlying outcomes space includes the use of lotteries and one assumes expected utility, the set $V(N)$ will be convex. These are the assumptions made by Nash (1950) in his description of a two-player bargaining problem, where $N = \{1, 2\}$.

- $d = (0, \dots, 0) \in V(N)$ is the disagreement utility point. In our notation, we have that $d_i = \max_{u_i \in V(\{i\})} u_i$. That is, d is the utility profile that will result from no agreement. The fact that one can take $d = 0$ is a normalization of all utility functions. This can be done without loss of generality if preferences can be represented equivalently by utility functions that differ by an additive constant.

- There are points in $V(N)$ that dominate d, i.e., there exists $u \in V(N)$ such that $u_i > d_i$ for all $i \in N$. This is a particular case of the property of superadditivity that will be defined later; in this case, it simply says that there are gains from cooperation in the grand coalition N.

Two major solutions to pure bargaining problems were proposed in Nash (1950) and in Kalai and Smorodinsky (1975), to which we devote the next subsections.

4.1 The Nash Bargaining Solution

Bilateral monopoly problems had been deemed indeterminate by earlier economic theorists, including Edgeworth and Hicks. Indeed, standard economic theory did not offer compelling enough arguments to single out a prediction from the set of agreements that are both individually rational (i.e., each party receiving a utility at least equal to her disagreement utility) and collectively rational (i.e., Pareto efficient). In non-trivial bargaining problems, the set of agreements with these properties is large.

Nash's breakthrough was to propose an approach with a sharp prediction. His idea was to employ the axiomatic method. Consider the class Γ of all two-player pure bargaining problems, in which the set $V(\{1,2\})$ is compact and convex. Consider single-valued solutions $f : \Gamma \mapsto \mathbb{R}^2$ such that $f(\{1,2\},V) \in V(\{1,2\})$ for every $(\{1,2\},V) \in \Gamma$. Nash (1950) imposes the following four properties on the solution f:

- Scale invariance: if one multiplies by arbitrary positive constants and adds arbitrary scalars to the utility functions of each player, the solution assigns to the transformed problem the same transformation of the utility pair originally assigned by the solution. This is justified when one works with von Neumann-Morgenstern expected utility.

- Symmetry: if the bargaining problem is symmetric so that $d_1 = d_2$ and $V(\{1,2\})$ is such that $(a,b) \in V(\{1,2\})$ if and only if $(b,a) \in V(\{1,2\})$, the solution assigns the same utility to both players. The solution must exhibit equal treatment of equals.

- Efficiency: the solution assigns a point on the Pareto frontier of $V(\{1,2\})$. It must exploit all available gains from cooperation.

- Independence of irrelevant alternatives (IIA): comparing two problems that share the same disagreement point, if $V(\{1,2\})$ is a subset of $V'(\{1,2\})$ and $f(\{1,2\},V') \in V(\{1,2\})$, then the solution must assign the same point to both problems. Nash justified this property by saying that there is "no action at a distance" in the determination of the solution.

Nash (1950) shows that, over the class of problems considered, there exists a unique solution that satisfies these four properties. It is the function that assigns to each bargaining problem the point that maximizes the product

$(u_1 - d_1)(u_2 - d_2)$ over all pairs $(u_1, u_2) \in V(\{1,2\})$. This function is what we refer to as the Nash solution, which has been used in many applications. It is worth pointing out that the particular functional form of the Nash solution, i.e., the point that maximizes the Nash product, far from being assumed, is an implication of the axioms. As it turns out, we will come back to rediscover the Nash product of utilities as playing an unexpected role in some of the implementation results.

A demand game: The first mechanism in the Nash program was proposed in Nash (1953); see Binmore (1987) for a nice presentation. Given a two-person bargaining problem in which the feasible set of utilities is $V(\{1,2\})$ and $d = 0$, each bargainer $i = 1, 2$ is asked to demand a utility level $\hat{u}_i \in \mathbb{R}_+$. These demands $(\hat{u}_1, \hat{u}_2)$ are submitted simultaneously. The outcome enforced is the one leading to the demand pair if it is feasible; otherwise, there is disagreement.

One can adapt this mechanism to the framework of implementation theory by formulating it in terms of physical outcomes. For example, suppose the underlying physical outcomes are the shares of a one-dimensional pie of size 1 if there is agreement, while no pie is created if there is disagreement. Then, the demands in the demand game can be formulated in physical terms. Physical demands are easier to enforce than are utility demands; according to this formulation, utility only appears in the background (the two bargainers will be evaluating outcomes according to their utility functions). We will proceed to the analysis of this mechanism.

Formally, in this mechanism, the set of strategies for an agent is the interval $[0, 1]$. Let x_1 and x_2 be the two physical demands made by players 1 and 2, respectively. The outcome (z_1, z_2) is the split (x_1, x_2) if $x_1 + x_2 \leq 1$; $(z_1, z_2) = (0, 0)$ otherwise. An outcome gives rise to a utility pair $(u_1(z_1), u_2(z_2))$, where the functions u_i are concave, strictly increasing, differentiable and normalized so that $u_i(0) = 0$.

The set of Nash equilibrium outcomes of this mechanism includes the entire Pareto frontier ((x_1^*, x_2^*) is a Nash equilibrium whenever $x_1^* + x_2^* = 1$), as well as disagreement ($(x_1^*, x_2^*) = (1, 1)$ with outcome $(z_1, z_2) = (0, 0)$ is also a Nash equilibrium). However, Nash (1953) proposed a refinement of the set of Nash equilibria based on the following perturbation of the demand game.

Suppose that bargainers do not know for sure the size of the pie they will be creating if they reach an agreement. Assume the size of the pie is a random variable τ, with density f, that includes $\tau = 1$ in its support. Now, given a pair of demands (x_1, x_2), the outcome is the pair of shares (x_1, x_2) if the pie realized makes them feasible, or disagreement. Specifically, given a demand pair, the expected payoff to player $i = 1, 2$ is

$$u_i(x_i)[1 - F(x_i + x_j)].$$

The first-order condition of expected payoff maximization for player $i = 1, 2$ is:

$$u_i'(x_i)[1 - F(x_i + x_j)] = u_i(x_i)f(x_i + x_j).$$

The interpretation of this equation is simple: the marginal utility gain attributed to a marginal increase in one's demand must be equated to the marginal utility loss due to the lower probability of the two demands being compatible. At a Nash equilibrium (x_1^*, x_2^*), the resulting sum of shares $x_1^* + x_2^*$ must be in the support of the random variable. Also, we can rewrite both first-order conditions as follows:

$$\frac{1 - F(x_1^* + x_2^*)}{f(x_1^* + x_2^*)} = \frac{u_1(x_1^*)}{u_1'(x_1^*)} = \frac{u_2(x_2^*)}{u_2'(x_2^*)}.$$

Note how the Nash product already shows up here: at the Nash equilibrium (x_1^*, x_2^*) the slope in the utility plane of the hyperbola $u_1 u_2 = u_1(x_1^*)u_2(x_2^*)$ is equated to the slope of the ex-post Pareto contour $u_1^{-1}(u_1) + u_2^{-1}(u_2) = x_1^* + x_2^*$. Thus, this tangency simply expresses the equality of marginal gains and losses for each player. Perhaps this is what Nash meant when, in motivating his IIA axiom, he was talking about "no action at a distance" in the determination of the bargaining solution. In other words, local considerations, especially the marginal change in the probability of an agreement being reached, lead to this tangency. It is worth noting that this feature is a consequence of the strategic analysis of the perturbed demand game, even before uncertainty is removed. Finally, it is now easy to see the following result, proved in Nash (1953).

Proposition 1 *If the distribution of possible pie sizes becomes degenerate around $\tau = 1$, the equilibrium payoff of the perturbed demand game converges to the Nash bargaining solution payoff of the original problem.*

We learn that in negotiations with uncertainty over the size of the pie, but where beliefs are concentrated around the true size, the Nash equilibrium logic in a demand game pushes bargainers to equate marginal utility gains and losses, and these marginal calculations yield an outcome that gravitates towards the Nash solution. See also Guth, Ritzberger and van Damme (2004) for a recent variant of the result.

An alternating offers bargaining game: Consider the procedure proposed in Rubinstein (1982). Players 1 and 2 are bargaining over how to split one dollar. Negotiations take place in discrete units of time: $t = 1, 2, \ldots$ Player 1 begins the game at time $t = 1$ by proposing an offer $(x, 1 - x)$ for $x \in [0, 1]$, where x is the share of the dollar that she is asking for herself, and $(1 - x)$ the share that she is offering to player 2. Next, player 2 accepts or rejects this offer. If he accepts, the offer is implemented and the game ends. If he rejects, one unit of time elapses and in period $t = 2$ player 2 will make a counteroffer $(y, 1-y)$ for $y \in [0, 1]$, to which player 1 must respond. If player 1 accepts this offer, it is implemented; if she rejects, she will make a new offer in the next period ($t = 3$), and so on. Players are impatient, and the common per period discount factor is $\delta \in (0, 1)$. Thus, if player i receives a share x_i in an agreement reached in period t, player i's utility is $\delta^{t-1} u_i(x_i)$, where we assume that u_i is concave and strictly increasing.[3] Perpetual disagreement has a payoff of 0 for both players.

Consider the following system of equations:

$$\begin{aligned} u_1(x_1) &= \delta u_1(x_1 + \alpha) \\ u_2(x_2) &= \delta u_2(x_2 + \alpha) \\ x_1 + x_2 + \alpha &= 1. \end{aligned}$$

The interpretation of the unknowns of this system are simple enough: x_i is a share that player i is offered as a responder, while α is the premium for being the proposer. By construction, each player is indifferent between being the responder and the proposer in the next period if the shares obey these equations. It is not hard to see that the system has a unique solution (x_1^*, x_2^*, α^*).

Furthermore, although the alternating offers game admits a large multiplicity of Nash equilibria, it constitutes one of the most striking successes of subgame perfect equilibrium (SPE) as a refinement of Nash equilibrium. Indeed, in the alternating offers game where player 1 is the first proposer, there is a unique SPE outcome, which is the immediate agreement on the

[3] An alternative interpretation of the model is one in which δ is the probability that negotiations continue after a rejection. For the subtle differences between the two specifications, the reader is referred to Binmore, Rubinstein and Wolinsky (1986) or Osborne and Rubinstein (1990, Chapter 4).

split $(x_1^* + \alpha^*, x_2^*)$. Similarly, if player 2 is the first to propose, the unique SPE outcome is the immediate agreement on $(x_1^*, x_2^* + \alpha^*)$; see Osborne and Rubinstein (1990, Chapter 3). This remarkable result also provides a sharp prediction to the bilateral bargaining problem. But unlike Nash's approaches, which rely on either the axioms or uncertainty in a simultaneous-move game, Rubinstein's uniqueness is a result of the use of credible threats in negotiations together with the assumption of impatience (or breakdown probabilities in the negotiations).

At the unique equilibrium, each player in the role of responder is offered a share that makes him/her exactly indifferent between accepting and rejecting the offer to continue bargaining in the next period. Therefore, the role of proposer is associated with a "premium," which we have called α^*. Moreover, the Nash product shows up again, perhaps in an unexpected way. Indeed, the two equilibrium splits corresponding to the games in which either player 1 or 2 are the first proposer, i.e., $(u_1(x_1^* + \alpha^*), u_2(x_2^*))$ and $(u_1(x_1^*), u_2(x_2^* + \alpha^*))$ generate the same Nash product. To see this, just use the first two equations in the above system. Again, it is remarkable that the "Nash product" property of the unique SPE payoffs prevails for any δ, even far away from the limit as $\delta \to 1$. Here, the constant value of the Nash product at the two equilibrium splits is an expression of equilibrium play under the use of credible threats.

The observation in the previous paragraph leads to the following result, shown in Binmore, Rubinstein and Wolinsky (1986) and Binmore (1987).

Proposition 2 *The limit as $\delta \to 1$ of the unique SPE payoff of the alternating offers game is the Nash bargaining solution payoff.*

One can see this result easily. The previous discussion has established that the two equilibrium splits (if either player 1 or player 2 are the first proposer) are characterized in the utility space by the intersections of the Pareto frontier of $V(\{1,2\})$ -both are agreements with no delay- and a symmetric hyperbola of equation $u_1 u_2 = K$. Moreover, as $\delta \to 1$, $\alpha^* \to 0$ (again, see the system of equations above), which implies that the two splits converge to the same. That split must correspond in utility space to the point where the Nash product is maximized over $V(\{1,2\})$.

Note how when δ is close to 1, the "no action at a distance" manifests itself in the equilibrium strategies of the game: equilibrium proposals and counterproposals occur in a small neighborhood of the Nash solution. This will contrast sharply with the implementation of the Kalai-Smorodinsky solution. That is, if bargainers envision themselves at the negotiation table and consider unilateral deviations from their equilibrium actions when δ is near 1, the kind of splits put on the table following those deviations are very near the final compromise.[4]

Extensions of this result to n-player pure bargaining problems are found in Krishna and Serrano (1996) and Chae and Yang (1994). Krishna and Serrano (1996) propose a procedure in which, while negotiations are multilateral –a proposal is an n-component split of the pie–, responders are given the option of signing a side-deal with the proposer and can exit the game by accepting. The rejectors, on the other hand, continue bargaining with the proposer, who is the legal representative of the acceptors in these further rounds. This exit feature proved crucial in obtaining uniqueness of the SPE prediction.[5] The flexibility of the bargaining procedure based on exit was inspired by the axiom of consistency, used by Lensberg (1988) in his characterization of the Nash solution. This was a good example of the usefulness of the axiomatic approach for the Nash program: whereas the individual veto power embodied in the unanimity extension facilitates the existence of a plethora of equilibria through complex schemes of reward and punishment in negotiations, the exit feature makes it impossible to punish a player for accepting a proposal. Following partial acceptances of a proposal, the negotiations can be "sliced" among fewer remaining parties and one can use mathematical induction and Rubinstein's uniqueness result to show uniqueness for any number of players. The procedure in Chae and Yang (1994), a variant of a game first analyzed in Jun (1987), is based on a sequence of bilateral meetings involving the proposer and one responder at a time. In a sense, the same flexibility of being able to slice the multilateral negotiations into smaller pieces is also key to their result.

Other mechanisms: All the implementation results of the Nash solution surveyed so far are approximate results. It is interesting to note other papers

[4] This will also be a feature of other more general models that, for the case of pure bargaining, yield the Nash solution (e.g., Hart and Mas-Colell (1996), Serrano (1997b)). This too happens in Herrero (1989), who analyzes the bargaining problem dropping convexity of the feasible set and finds how the equilibrium payoffs converge to tangency points between the Nash product hyperbolas and the Pareto frontier.

[5] The large multiplicity of SPE payoffs in the "unanimity" extension of the two-person alternating offers game, in which each responder has veto power, are well described in Osborne and Rubinstein (1990, chapter 3).

that yield the Nash solution exactly as the result of equilibrium play. In an approach strongly rooted in standard implementation theory, Howard (1992) implements the Nash solution in SPE using a multi-stage sequential mechanism; see also Miyagawa (2002), who proposes sequential mechanisms to implement a class of bargaining solutions that maximize an increasing function of utilities, among them the Nash solution. Trockel (2002c) offers another exact implementation of the Nash solution, in which the message sets contain bargaining solutions (agents are "bargaining over solutions"); see also Naeve (1999). In Trockel (2000) a "Walrasian implementation" of the Nash solution is proposed: the mechanism uses relative utility prices given by the normal tangents to the frontier of the feasible set $V(\{1,2\})$ of the bargaining problem.

4.2 The Kalai-Smorodinsky Bargaining Solution

As an alternative to the Nash solution, Kalai and Smorodinsky (1975) propose a different way to solve bargaining problems. They impose the first three axioms of Nash's characterization (scale invariance, symmetry and efficiency), but IIA is dropped. Instead, a different axiom is proposed. To state it, we should first introduce a previous definition.

Given a two-person bargaining problem in which the disagreement utilities have been normalized to 0, we call player i's aspiration level $a_i(\cdot)$ to the maximum feasible utility that player $i$can achieve if player $j \neq i$ receives a utility of 0. That is, the highest utility player i may aspire to in the bargaining problem is the one corresponding to an agreement where she extracts all the surplus, her dictatorship outcome. Now we are ready to state the new axiom for a bargaining solution f:

- Individual monotonicity: if two bargaining problems share the zero disagreement point and if $V(\{1,2\})$ is a subset of $V'(\{1,2\})$ but $a_j(V) = a_j(V')$, the solution must assign a utility to player $i \neq j$ in the bigger problem that is at least as high as in the smaller problem.

Unlike IIA, which incorporates local information around the solution into the derivation of the settlement, individual monotonicity emphasizes the importance of "global" considerations as expressed in the aspiration levels. Paraphrasing Nash's words, this axiom is a way to take "action at a distance" into account in order to get to the final settlement. Extreme threats based on the utility derived from dictatorship may be heard now.

Given a bargaining problem, let $a = (a_1, a_2)$ be the point whose coordinates are the aspirations of both players. This point will typically be not feasible, but one can scale utilities down proportionally. The Kalai-Smorodinsky solution is the function that assigns to each bargaining problem the intersection point of the Pareto frontier of $V(\{1, 2\})$ with the straight-line segment $0 - a$. It is the only solution satisfying scale invariance, symmetry, efficiency and individual monotonicity over the class of compact and convex bargaining problems. Note the major role played by the two dictatorship outcomes in the determination of the solution.

Bidding for fractions of dictatorship: Moulin (1984) proposes a procedure that exactly implements in SPE the Kalai-Smorodinsky solution for any number of players. For simplicity in the presentation, we shall describe it only for the bilateral case. Recall that one can normalize the disagreement utilities to 0.

The mechanism begins with both players making a simultaneous announcement, a bid $p_k \in [0, 1]$ for $k = 1, 2$. The player whose bid is the highest will make a proposal to the other (if there is a tie, any predetermined tie-breaking rule will choose the proposer –say, player 1). Let player i be the proposer. Then, player i makes a proposal to player j. If j accepts, the proposal is implemented; otherwise, player j makes a "take it or leave it" offer to player i with probability p_i, while disagreement happens with probability $(1 - p_i)$. Thus, the bids of the first stage serve to determine the proposer, but also the winning bid is used as the fraction of dictatorship assigned to the non-winning bidder in the event of a rejection. The two forces create a tradeoff in the size of the bid one should employ. Also, note the prominent use in the mechanism of subgames in which there is "action at a distance" (i.e., subgames resolved through dictatorship by making use of the aspiration levels).

Moulin (1984) shows the following result:

Proposition 3 *The unique SPE payoff of the bidding game on fractions of dictatorship is the Kalai-Smorodinsky solution payoff of the bargaining problem.*

It is easy to see that, given a bid p_2 used by player 2, player 1 can guarantee at least the payoff corresponding to the efficient point in which player 2 receives a utility equal to $p_2 a_2$ – player 1 can do this simply by naming $p_1 = p_2$ and making that proposal, which will be accepted by player

2. Let us call this payoff eff$(p_2 a_2)$. In addition, player 1 can approximate a payoff of $p_2 a_1$ simply by bidding $p_1 = p_2 - \epsilon$ for any arbitrarily small $\epsilon > 0$. Thus, player 1 can guarantee the $\max\{\text{eff}(p_2 a_2), p_2 a_1\}$, and therefore, player 2, in trying to minimize player 1's payoff (recall that all these outcomes are efficient) will choose p_2 so as to equate the two terms in the maximum expression, i.e., eff$(p_2 a_2) = p_2 a_1$. But these arguments imply that, while player 1's equilibrium payoff is $p_2 a_1$, player 2's is $p_2 a_2$, so their ratio is the ratio of aspiration levels. Furthermore, the equilibrium is efficient, so the outcome is the Kalai-Smorodinsky solution.

Once again note how this implementation relies heavily on the extreme threats embodied in the dictatorship outcomes ("action at a distance"). According to the equilibrium strategies, if the proposal is rejected by the responder, she would become a dictator in the "take it or leave it" subgame. As we have seen in the previous subsection, this extreme behavior at the negotiation table is not part of the equilibrium in the procedures that implement the Nash solution. See Trockel (1999) and Haake (2000) for other implementations of the Kalai-Smorodinsky solution.

4.3 Two Examples

For a comparison between the implementations of the Nash and the Kalai-Smorodinsky solutions, consider the following examples, simply meant to illustrate some of the points already made.

Example 1 *Both bargainers are trying to reach an agreement in order to split one dollar. Let the utility function of player 1 be linear in money:*

$$u_1(x_1) = x_1.$$

However, player 2's utility function exhibits a drop in her marginal utility when she is awarded at least one half of the surplus:

$$u_2(x_2) = \begin{cases} x_2 & \textit{if} \quad x_2 \leq 1/2; \\ 0.5 + \epsilon(x_2 - 0.5) & \textit{otherwise}; \qquad \epsilon \in (0,1) \end{cases}$$

It can be shown that the equilibrium splits in the Rubinstein game are

$$(x_1^* + \alpha^*, x_2^*) = \Big(\frac{1 - 0.5\delta(1+\epsilon)}{1 - \delta^2\epsilon}, \frac{0.5\delta(1+\epsilon) - \delta^2\epsilon}{1 - \delta^2\epsilon}\Big)$$

if player 1 is the first to propose, and

$$(x_1^*, x_2^* + \alpha^*) = \big(\frac{\delta(1 - 0.5\delta(1+\epsilon))}{1-\delta^2\epsilon}, \frac{1 - \delta - 0.5\delta^2\epsilon + 0.5\delta^2}{1-\delta^2\epsilon}\big)$$

if player 2 begins the game. The proposer's premium, equal for both players, is $\alpha^* = \frac{(1-\delta)(1-0.5\delta(1+\epsilon))}{1-\delta^2\epsilon}$*, which goes to 0 as* $\delta \to 1$*. Note how as* $\delta \to 1$*, both equilibrium splits approximate the equal split of the pie, regardless of* ϵ*. In the differentiable version of this problem, the only point of the Pareto frontier of the bargaining problem in which the utility elasticity is 1 continues to be* $(0.5, 0.5)$ *for any value of* ϵ*, and the Rubinstein equilibrium splits, dictated by the equality of the Nash product, gravitate towards that point.*

In contrast, the unique SPE payoff in the Moulin game is

$$(u_1^*, u_2^*) = \big(\frac{1}{1+0.5(1+\epsilon)}, \frac{0.5(1+\epsilon)}{1+0.5(1+\epsilon)}\big),$$

which is the Kalai-Smorodinsky solution payoff. This utility pair, which coincides with the shares of the dollar, is of course sensitive to ϵ*: the larger the drop in 2's marginal utility the lower her aspiration level, which reflects itself in lower equilibrium shares. Moulin's equilibrium continues to use the extreme threats in the dictatorship subgames, but 2's dictatorship threat becomes less effective when her marginal utility for getting the entire surplus is lower.*

Example 2 *Consider again the split of a dollar between two players. Again player 1's utility function is*

$$u_1(x_1) = x_1,$$

whereas player 2's is

$$u_2(x_2) = x_2^\epsilon, \qquad \epsilon \in (0,1).$$

For any ϵ*, no change in the aspiration level of player 2's takes place. Because of this, in utility terms, the unique SPE payoff of the Moulin game yields a point on the 45 degree line of the positive orthant,* $u_1 = u_2$*. Of course, risk aversion hurts player 2 when* ϵ *goes down: her equilibrium share of the dollar is* $0.5^{\frac{1}{\epsilon}}$*.*

However, the point on the Pareto frontier in which $u_1 = u_2$ *does not have a unit utility elasticity. It can be checked that the "Nash product property"* *of the Rubinstein equilibrium splits places them below the diagonal of the orthant when δ is large enough, at points in which $u_1 < u_2$: in response to the change in local conditions, player 2 is less well treated by this bargaining (when compared to the Moulin procedure).*

5 Coalitional Games

Now we shall consider the class of general coalitional games (N, V) in which subcoalitions have power. Examples include the coalition formation facing the different political parties in the parliament, negotiations among countries in an international context, or the disputes among creditors over the estate of a bankrupt firm. Recall that $V(S)$ represents the set of feasible utility profiles for coalition S. An important subclass of coalitional games is that of transferable-utility (TU) games, in which the sets $V(S)$ are half-spaces whose Pareto frontier is the hyperplane of equation $\sum_{i \in S} x_i = v(S)$, where the real number $v(S)$ is referred to as the worth of S. Abusing notation slightly, we shall refer to the pair (N, v) as a TU game, where v is the characteristic function that assigns the number $v(S)$ to each coalition S; without loss of generality, we assign 0 as the worth of the empty coalition, and we can also normalize to 0 the worth of each individual player: $v(\{\emptyset\}) = 0$ and $v(\{i\}) = 0$ for every $i \in N$. TU games are a representation of a situation in which utility is transferable at a one-to-one rate among players, for example, because utility is expressed in units of a nummeraire.

Some properties of TU games that are often used are the following. We list them from weaker to stronger:

- Monotonicity: larger coalitions are worth more, i.e., whenever $S \subseteq T$, $v(S) \leq v(T)$.
- Superadditivity: not only is the worth of a larger group larger, but it exceeds the sum of its parts. That is, any two disjoint subcoalitions always find it profitable to cooperate: whenever $S \cap T = \emptyset$, $v(S) + v(T) \leq v(S \cup T)$.
- Convexity: not only are there gains from cooperation, but those are increasing in the size of the groups, i.e., for every $S \subseteq T$ and every $i \notin T$, $v(S \cup \{i\}) - v(S) \leq v(T \cup \{i\}) - v(T)$.

In this section we shall focus mainly on the two major single-valued solutions that have been proposed for TU games: the Shapley value and the nucleolus. We will deal more briefly with set-valued solutions in a final subsection.

5.1 The Shapley Value of a TU Game

We are interested in single-valued solutions f. That is, for each TU game (N, v), the question posed now is whether one can always find a single payoff vector $f(N, v)$ that will represent a "fair" retribution to each player, when one takes into account the worths of all coalitions. Shapley (1953) imposes the following four properties on a solution f:

- Efficiency: in the sense of budget-balance for the grand coalition, i.e., for all games (N, v), $\sum_{i \in N} f_i(N, v) = v(N)$.
- Symmetry: two players that are identical in terms of their contributions to all coalitions should be paid the same by the solution.
- Additivity: if one defines a sum game $(N, [v + w])$, where for each S $[v + w](S) = v(S) + w(S)$, we impose that $f(N, [v + w]) = f(N, v) + f(N, w)$.
- Dummy: if a player contributes 0 to each coalition that she joins, the solution should pay her 0.

Shapley (1953) shows that, over the entire class of TU games, there exists a unique function satisfying these four properties. It is the function $\mathrm{Sh}(N, v)$ that assigns to each player i in the game (N, v) the amount

$$\mathrm{Sh}_i(N, v) = \sum_{S, i \in S} \frac{(|S| - 1)!(|N| - |S|)!}{|N|!} [v(S) - v(S \setminus \{i\})].$$

That is, the Shapley value awards to each player the average of her marginal contributions to each coalition. In taking this average, all orders of the players are considered to be equally likely.

In Young (1985), a marginality axiom serves to characterize the Shapley value, together with efficiency and symmetry – Shapley's and Young's theorems are to be viewed in parallel: while the former derives the marginal

formula from additivity, the latter obtains the linear dependence on the marginal contributions from the marginality axiom. Hart and Mas-Colell (1989) arrive at the Shapley value from the concept of a potential function – another instance of marginality, since the Shapley value happens to be the gradient of the unique potential function associated with a game. The content of these axioms is present in the different implementation results that we survey now.
"If my offer is rejected today, I may not be around tomorrow": Based on a simpler procedure studied in Mas-Colell (1988), Hart and Mas-Colell (1996) propose the following mechanism. At every round, there is a set of active players and the proposer is chosen at random among them with equal probability (in the first round all players in N are active). Once a player is appointed as the proposer, she makes a feasible proposal to the set of active players. Responses are sequential according to some specified order. If all accept the proposal, it is implemented. If at least one responder rejects it, a random device is used. While with probability $\rho \in [0, 1)$ the set of active players is unchanged and a new round starts, with probability $(1 - \rho)$ the proposer is removed from the set of active players: she is left to use her individual resources, and the others become the new set of active players in a new round.

The result shown in Hart and Mas-Colell (1996) for TU games is the following:

Proposition 4 *Let (N, v) satisfy monotonicity. For any ρ, the unique stationary SPE payoff of the bargaining game with random removal of proposers is the Shapley value payoff.*

The restriction to stationary strategies is not new; in fact, it was already explored in certain multilateral extensions of the Rubinstein game. In the Hart and Mas-Colell game, stationarity amounts to the proposals taking into account only the current active set of players, and responses being made in terms of the payoff offered to a player in the given offer, thereby disregarding other aspects of history.[6]

The intuition is perhaps most clear when $\rho = 0$. Start analyzing a subgame where the set of active players consists of only two players. Clearly, the proposer can extract in the offer that she makes her marginal contribution

[6] Krishna and Serrano (1995) remove the stationarity assumption and show that the uniqueness result remains as long as $\rho \leq 1/(|N| - 1)$. This is reminiscent of a result in Herrero (1985) for the pure bargaining unanimity game.

to the other player. Since the proposer is chosen at random, the unique SPE of this bilateral subgame yields the Shapley value payoff. By induction, suppose this much is true for subgames consisting of $(|N|-1)$ players. Since the proposer and her resources will be removed if her offer is rejected, she can extract in it her marginal contribution to the remaining coalition. Again, the random choice of proposers ensures that the final payoffs are the ones given by the Shapley value of the game. The same intuition carries over when $\rho \in (0, 1)$.

Note a few salient features of this procedure. First of all, the presence of randomness is important, both in the choice of proposers and in the determination of the active set of players. Randomness is usually defended as a way to preserve anonymity in the procedure. Second, if players drop out, their resources go with them. As outlined above, this is the way in which marginal contributions of players to coalitions become crucial. And third, rejected proposals cause players to drop out, but the proposal itself is not essential to the ensuing outcome: it is not used at all in the continuation. As such, one could say that it is the worth of coalitions in the characteristic function, and not the specifics of a given proposal, what determines the final outcome of negotiations. When an offer is rejected and the proposer is removed as an active player, the payments that she proposed and the interaction of her resources with those of others are completely ignored for the rest of the bargain.

These features are also found in other implementation results of the Shapley value. Pérez-Castrillo and Wettstein (2001) study a procedure similar to that of Hart and Mas-Colell's, but in which the random choice of proposers is replaced with a bidding stage. Something similar happens in Winter (1994a), who obtains an implementation result for the Shapley value of convex games. The game in Gul (1989) differs from the ones just described mainly in that the encounters at the negotiation table are always bilateral: if a proposal is accepted, the acceptor "sells" her resources to the proposer and negotiations continue in which the player who has "bought off" the resources of other players effectively controls them. There is discounting across periods. Gul's result is that every stationary SPE that exhibits immediate agreements yields payoffs that are close to the Shapley value payoffs when discounting is removed (see also Hart and Levy (1999) and Gul (1999)).

One important direction for the Nash program is also suggested in Hart and Mas-Colell (1996). Namely, the attempt is to apply the same extensive form to larger domains of coalitional games beyond the TU class. This can potentially shed light on two questions: (a) to establish connections among already known solution concepts that arise via the bargaining procedure, and (b) to uncover extensions of solution concepts to domains where the theory is less settled. Indeed, as already indicated in the previous section, they showed that as $\rho \to 1$ the average stationary SPE payoff converges to the Nash solution of a pure bargaining problem. Furthermore, when the underlying environment is a general non-transferable (NTU) game, the stationary SPE payoffs of the procedure converge to the payoffs prescribed by the consistent value of Maschler and Owen (1989, 1992).

Following this agenda, but in a different procedure, de Clippel (2004) also obtains the Shapley value as its unique SPE payoff over the class of TU games. However, his procedure yields a different solution when applied to pure bargaining problems and a new value concept over the class of NTU games: in bargaining problems, de Clippel's procedure converges to the Raiffa solution –more similar to Kalai-Smorodinsky's–, and in general NTU games to what he calls the "procedural" value, different from the other known extensions of the Shapley value. In the de Clippel mechanism, bargaining lasts R rounds for each given set of active players. In round $r < R$ the Hart and Mas-Colell procedure is played for the case of $\rho = 1$, whereas at round R the $\rho = 0$ version of the Hart and Mas-Colell game is activated. In words, when $r < R$ a rejection is followed by a new random choice of proposer among the current set of active players, while at round R the rejected proposer and her resources are removed from the game, and the game continues with the new set of active players. The result is obtained as $R \to \infty$.

Finally, in a similar endeavor, but this time considering extensions to games in partition function form, Maskin (2003) studies a procedure that yields the Shapley value over some classes of characteristic functions, and a new solution is proposed as the SPE outcome of the same game in the case of partition functions that allow externalities.

5.2 The Nucleolus of a TU Game

A different single-valued solution concept for TU games is the nucleolus. Suppose one is interested in the notion of coalitional stability as formalized by the core, but wishes to insist on a solution that: (a) is always non-empty; and (b) prescribes a single payoff vector for each game. The nucleolus is one such core selection. The nucleolus was characterized by making use of different axioms.[7]

- Covariance: if one multiplies by the same arbitrary positive constant and adds arbitrary scalars to the utility functions of each player, the solution assigns to the transformed problem the same transformation of the utility profile originally assigned by the solution. This is justified when one works with von Neumann-Morgenstern expected utility and one wishes to remain in the TU class after the transformation.

- Consistency: the predictions of the solution should be invariant to the number of players in the game. Specifically, given a payoff vector $x \in \mathbb{R}^N$ and a game (N, v), one can define the Davis-Maschler (DM) reduced game for coalition $S \subseteq N$, (S, v_{xS}), as follows: $v_{xS}(S) = v(N) - \sum_{j \notin S} x_j$, while for non-empty proper subsets T of S, $v_{xS}(T) = \max_{Q \subseteq N \setminus S}\{v(T \cup Q) - \sum_{j \in Q} x_j\}$. Then, for every (N, v), every $S \subseteq N$ and every $i \in S$, f is consistent whenever $f_i(N, v) = f_i(S, v_{xS})$.

A brief explanation regarding consistency is in order (see Thomson (1990)). The idea is that the solution should be invariant to any subset of players leaving the room and, after making use of some interactions between their resources and those of the rest of the players, trying to apply the same solution. The DM reduced game, first proposed in Davis and Maschler (1965), presupposes a payoff vector that is tentatively considered as the final payoff. This reduced game stipulates that while the entire group that leaves the room (coalition S) is left with $v(N) - \sum_{j \notin S} x_j$ for feasibility reasons, subcoalitions T of S can threaten each other to determine their "virtual bargaining power" by choosing optimally the subset of $N \setminus S$ with whom they wish to cooperate. It is important to note that, apart from the nucleolus, the core and the kernel also satisfy consistency in this sense.

In the first use of a consistency axiom to characterize a game theoretic solution, Sobolev (1975) characterizes the nucleolus as the unique single-valued solution satisfying symmetry, covariance and consistency (see Peleg and Sudholter (2003, Chapter 6)). For completeness, we provide at present the definition of the nucleolus (Schmeidler (1969)).

[7] Here I am using the term "nucleolus" to refer to what is sometimes termed the "prenucleolus." The difference between the two is that the former imposes individual rationality, while the latter does not. In superadditive games, the ones we shall discuss, the prenucleolus is individually rational, so both prescribe the same payoff.

For every vector $x \in \mathbb{R}^N$, and for every coalition $S \subseteq N$, we define a number $e_S(x)$ to be the excess of coalition S at the payoff vector x, that is, $e_S(x) = \sum_{i \in S} x_i - v(S)$. We interpret this excess as a measure of welfare of coalition S if the proposed vector is x. The bigger is $e_S(x)$, the bigger the welfare of coalition S at x (that is, the less S would gain by departing from x). Suppose we have a preestablished order for the coalitions $S_1, \ldots, S_{2^n}$. Then, we call $e(x)$ the vector of all excesses, that is, $e(x) \in \mathbb{R}^{2^n}$, with generic element $e_i(x) = \sum_{j \in S_i} x_j - v(S_i)$, $i = 1, \ldots, 2^n$. The nucleolus of the game (N, v) is the set of payoff vectors $x \in \mathbb{R}^N$ such that $\sum_{i \in N} x_i = v(N)$ and satisfying that there does not exist y, $\sum_{i \in N} y_i = v(N)$ such that $e(y)$ is leximin superior to $e(x)$. (A vector is leximin superior to another if, when the coordinates of both are arranged in increasing order, the former is lexicographic superior to the latter). It turns out that the set of vectors satisfying this definition is always a singleton.

"If my offer is rejected today, renegotiations with the rejectors are called for": In the context of bankruptcy problems, Aumann and Maschler (1985) find an amazing application of the nucleolus. Given a bankruptcy problem (E, d), where $E \geq 0$ is an estate to be divided among n creditors, and $d_i \geq 0$ is creditor i's claim, $E \leq \sum_{i \in N} d_i$, one can define a game in characteristic function (N, v). In it, N is the set of creditors, and for every coalition S, $v(S) = \max\{0, E - \sum_{i \notin S} d_i\}$. Then, as argued in Aumann and Maschler (1985), the nucleolus of this game coincides with the splits of the estate recommended in the Jewish Talmud to solve bankruptcy problems. In particular, the nucleolus of a two-player game so constructed is the split given by the Talmudic "contested garment" (CG) rule to a two-creditor problem:

$$\mathrm{CG}_i(E, (d_i, d_j)) = \max\{0, E - d_j\} + \frac{E - \max\{0, E - d_i\} - \max\{0, E - d_j\}}{2}.$$

For bankruptcy problems, Serrano (1995a) proposes the following bargaining procedure. Choose one of the creditors with the highest claim to make an offer to the other creditors. Let the proposer be creditor 1 and her offer be $x = (x_1, \ldots, x_n)$. Responses are sequential according to some arbitrary order. If creditor j accepts the offer, she receives x_j from the proposer and leaves the game. Let A be the set of acceptors of the proposal. If creditor j is the first to reject the proposal according to the order, she receives $z_j \mathrm{CG}_j(x_1 + x_j, (d_1, d_j))$ and creditor 1's share is modified to $z_1^j = x_1 + x_j - z_j$. After this renegotiation with j, creditor 1 takes up the other rejectors in similar bilateral renegotiations. Say the next one is creditor k. Then, creditor k receives $z_k = \mathrm{CG}_k(z_1^j + x_k, (d_1, d_k))$ and 1's share is adjusted to $z_1^k = z_1^j + x_k - z_k$, and so on. Creditor 1's final payoff is her share after the bilateral renegotiation with the last rejector.

Proposition 5 *For bankruptcy problems, in the bargaining procedure with bilateral renegotiation appeals, if the proposer is one of the creditors with the highest claim, the unique SPE payoff is the nucleolus of the associated characteristic function game.*

It is instructive to compare the salient features of this procedure to those outlined for the Shapley value implementations. First, the procedure allows for "partial agreements," inspired by consistency. Implicitly, the proposer is given more power, since she can release a creditor from further negotiations simply by getting her to accept the offer. Second, when a rejection takes place, the proposal itself is essential to the ensuing outcome: it is used to determine the estate of each bilateral renegotiation, in which a "fair" allocation between proposer and rejector is sought. Finally, the proposer must be chosen carefully. Far from being chosen at random, only when the proposer is one with the highest claimant does uniqueness of the SPE obtain. It is interesting to see how an important feature of the solution seems to show up in the way one must choose proposers: while the randomness of the Shapley value manifests itself in the random choice of proposers, in the nucleolus, an outcome of maximization of some sort, the identity of the proposer is also the solution to some "maximization" problem. Interestingly, Aumann and Maschler (1985, Section 5) describe a coalitional procedure that also yields the nucleolus: in this procedure, the only coalitions that form are the set of holders of the $(n-1)$ highest claims and the lowest claimant (different coalitions in that procedure would not yield the nucleolus shares).

Serrano (1993) provides a related implementation result of the nucleolus of all 3-player superadditive games. Developing a variant of a model in Okada (1996), Yan (2002) supports in stationary SPE a single-valued core selection different from the nucleolus. Vidal-Puga (2004) proposes a backward induction implementation of a new single-valued solution concept, which he terms the "selective value." The selective value incorporates aspects of the DM reduced game, which as already noted is of importance for the nucleolus. Finally, Sonmez (1999) provides a remarkable result relating single-valued cores and implementation in dominant strategies.

5.3 An Example

As we already saw in the pure bargaining section, it will be instructive to look at an example as a way to draw a comparison between the implementations of the Shapley value and the nucleolus:

Example 3 *Consider a 3-player game* (N, v)*, in which* $v(N) = 60$*,* $v(\{1, 2\}) = v(\{1, 3\}) = 54$*, while* $v(S) = 0$ *for all other coalitions* S*.*

The reader can calculate that the Shapley value awards the split $(38, 11, 11)$*, obtained in the Hart and Mas-Colell's (1996) game when* $\rho = 0$ *as the average of the following three vectors:* $(60, 0, 0)$ *made when player 1 is the proposer,* $(27, 6, 27)$ *when it is player 2, and* $(27, 27, 6)$ *when player 3 makes the offer. Note the prevalence of the logic of marginal contributions to coalitions. When player 1 proposes, she is aware that a rejection of her proposal would destroy all possibilities of positive surplus between 2 and 3, and hence, she can get away with a proposal where she extracts all the joint surplus. On the other hand, if either player 2 or 3 propose, a rejection would lead to a symmetric subgame between the two active players, who would have to split a surplus of 54; this is why each responder must be offered 27. In this example, the Shapley value is not in the core (coalitions* $\{1, 2\}$ *and* $\{1, 3\}$ *improve upon it) because the proposals made by either player 2 or 3 take much power away from player 1 in favor of the other responder.*

On the other hand, the nucleolus awards the shares $(54, 3, 3)$*. Expressing the game in terms of a bankruptcy problem, the estate to be shared is 60 and the vector of claims is* $(60, 6, 6)$*. The nucleolus is obtained in Serrano's (1995a) procedure when the power of making the proposal is given to player 1 (this will avoid the power loss that she suffered in the Shapley value procedure). In thinking what offer to make, creditor 1 may reason as follows: "well, the offer will be* (x_1, x_2, x_3)*. Creditors 2 and 3 are symmetric, so I'll offer* x *to each, and* $x = x_2 = x_3$ *should satisfy the bilateral renegotiation equation* $x = CG_i(x_1 + x, (60, 6))$ *for* $i = 2, 3$*. The solution to this is* $x = 3$*." The unique SPE outcome is the nucleolus of the associated characteristic function game, which of course is in the core. The reader can check that making this offer maximizes creditor 1's payoff, if she assumes that creditors 2 and 3 respond optimally to her proposal. To establish this step, it is important that the nucleolus of games generated by bankruptcy problems is monotonic in the estate. However, the nucleolus is not monotonic in general, and this fact has created difficulties for its implementation by means of* *consistency-based mechanisms.*

5.4 Set-Valued Solution Concepts

In this final subsection, we mention contributions to the Nash program concerning set-valued solution concepts. We briefly outline the insight achieved by the non-cooperative analysis into each of these solutions.

The core: Together with the Shapley value, the core is probably the most popular of the different cooperative solution concepts. This has translated in a large number of contributions to its non-cooperative implementation. The message that emerges from these different mechanisms that implement the core is that coalitions can be put together by a leader who takes the initiative of making proposals. To get the entire core as the set of equilibrium payoffs, one needs to restore some "anonymity" to the mechanism, neutralizing the advantage of the first mover. This has been accomplished through randomness in the choice of proposer, stationarity in the solution concept, or by making use of other devices by which proposals are not made by anyone in the set of players in the game (they are made by either extra players or are exogenously found as status quo situations). A partial list of these and related contributions includes the following papers: Alcalde, Pérez-Castrillo and Romero-Medina (1998), Banks and Duggan (2000), Baron and Ferejohn (1989), Bergin and Duggan (1999), Bloch (1996), Chatterjee et al. (1993), Evans (1997), Kalai, Postlewaite and Roberts (1979), Lagunoff (1994), Moldovanu and Winter (1994, 1995), Montero (1999), Okada (1996), Pérez-Castrillo (1994), Perry and Reny (1994), Selten (1981), Selten and Wooders (1991), Serrano (1995b), Serrano and Vohra (1997), Wilson (1978), Winter (1994b, 1996), Young (1998).

The bargaining sets: The different definitions of bargaining sets are an attempt to model the credibility of the objections raised by coalitions (see Aumann and Maschler (1964), Davis and Maschler (1967), Mas-Colell (1989), Dutta et al. (1989)). From the point of view of implementation, the objection/counterobjection infeasibility implicit in these definitions can be solved via sequentiality of moves. A useful trick to this effect, proposed in Serrano and Vohra (2002a, b), is for the proposer to have to ratify a coalitional agreement after she makes the proposal and the rest of the coalition accepts it. Other papers get around this difficulty by either allowing infeasible off-equilibrium outcomes (Einy and Wettstein (1999)) or extra players in the mechanism (Pérez-Castrillo and Wettstein (2000)); see also Morelli and

Montero (2003) for a different definition of the bargaining set. All these contributions help to clarify these different versions of bargaining sets, but more work is needed to better capture notions of credible objections in non-cooperative coalitional bargaining.

The kernel: The kernel was introduced in Davis and Maschler (1965) and axiomatized by Peleg (1986) using the DM reduced game. Inspired by this, Serrano (1997) proposes a non-cooperative model that yields the kernel in stationary equilibria. The way bilateral negotiations are portrayed between any pair of players generalizes the ideas expressed for the implementation of the nucleolus. In addition, the kernel is reinterpreted, shown to be free from cardinal interpersonal utility comparisons. In Montero (2002), a non-cooperative bargaining model is developed to support the kernel of apex games.

The stable sets: The stable sets were proposed by von Neumann and Morgenstern (1944). One problem with the definition, which disregards indirect dominance, was highlighted in Harsanyi (1974), who arrived at this criticism as a consequence of his non-cooperative analysis.

References

Alcalde, J., D. Pérez-Castrillo and A. Romero-Medina (1998), "Hiring Procedures to Implement Stable Allocations," *Journal of Economic Theory* **82**, 469-480.

Aumann, R. J. (1987), "Game Theory," in J. Eatwell, M. Milgate and P. Newman (eds.) *The New Palgrave Dictionary of Economics*, Norton, New York.

Aumann, R. J. (1997), "On the State of the Art in Game Theory," an interview, in *Understanding Strategic Interaction, Essays in Honor of Reinhard Selten*, W. Albers et al. (ed.), Springer, Berlin, 8-34.

Aumann, R. J. and M. Maschler (1964), "The Bargaining Set for Cooperative Games," in M. Dresher, L. S. Shapley and A. W. Tucker (eds.) *Advances in Game Theory*, Princeton University Press, 443-476.

Aumann, R. J. and M. Maschler (1985), "Game Theoretic Analysis of a Bankruptcy Problem from the Talmud," *Journal of Economic Theory* **36**, 195-213.

Banks, J. and J. Duggan (2000), "A Bargaining Model of Collective Choice," *American Political Science Review* **94**, 73-88.

Baron, D. and J. Ferejohn (1989), "Bargaining in Legislature," *American Political Science Review* **83**, 1181-1206.

Bergin, J. and J. Duggan (1999), "An Implementation-Theoretic Approach to Non-Cooperative Foundations," *Journal of Economic Theory* **86**, 50-76.

Binmore, K. (1987), "Nash Bargaining Theory (II)," Chapter 4 in K. Binmore and P. Dasgupta (ed.) *The Economics of Bargaining*, Basil Blackwell, Cambridge.

Binmore, K., A. Rubinstein and A. Wolinsky (1986), "The Nash Bargaining Solution in Economic Modelling," *RAND Journal of Economics* **17**, 176-188.

Bloch, F. (1996), "Sequential Formation of Coalitions with Fixed Payoff Division," *Games and Economic Behavior* **14**, 90-123.

Bossert, W. and G. Tan (1995), "An Arbitration Game and the Egalitarian Bargaining Solution," *Social Choice and Welfare* **12**, 29-41.

Chae, S. and J.-A. Yang (1994), "An n-Person Pure Bargaining Game," *Journal of Economic Theory* **62**, 86-102.

Chatterjee, K., B. Dutta, D. Ray and K. Sangupta (1993), "A Non-Cooperative Theory of Coalitional Bargaining," *Review of Economic Studies* **60**, 463-477.

Corchón, L. (1996), *The Theory of Implementation of Socially Optimal Decisions in Economics*, St. Martin's Press, New York.

Dagan, N. and R. Serrano (1998), "Invariance and Randomness in the Nash Program for Coalitional Games," *Economics Letters* **58**, 43-49.

Davis, M. and M. Maschler (1965), "The Kernel of a Cooperative Game," *Naval Research Logistics Quarterly* **12**, 223-259.

Davis, M. and M. Maschler (1967), "Existence of Stable Payoff Configurations for Cooperative Games," in M. Shubik (ed.) *Essays in Mathematical Economics* *in Honor of Oskar Morgenstern*, Princeton University Press, 39-52.

de Clippel, G. (2004), "The Procedural Value for Cooperative Games with Non-Transferable Utility," Mimeo, Brown University, forthcoming in *Mathematical Social Sciences*.

Dutta, B., D. Ray, K. Sengupta and R. Vohra (1989), "A Consistent Bargaining Set," *Journal of Economic Theory* **49**, 93-112.

Einy, E. and D. Wettstein (1999), "A non-cooperative interpretation of bargaining sets," *Review of Economic Design* **4**, 219-230.

Evans, R. (1997), "Coalitional Bargaining with Competition to Make Offers," *Games and Economic Behavior* **19**, 211-220.

Gul, F. (1989), "Bargaining Foundations of the Shapley Value," *Econometrica* **57**, 81-95.

Gul, F. (1999), "Efficiency and Immediate Agreement: a Reply to Hart and Levy," *Econometrica* **67**, 913-917.

Guth, W., K. Ritzberger and E. van Damme (2004), "On the Nash Bargaining Solution with Noise," *European Economic Review* **48**, 697-713.

Haake, C.-J. (2000), "Support and Implementation of the Kalai-Smorodinsky Bargaining Solution," in K. Inderfurth et al. (ed.) *Operations Research Proceedings 1999*, Springer Verlag, pp. 170-175.

Harsanyi, J. C. (1974), "An Equilibrium Point Interpretation of Stable Sets and a Proposed Alternative Definition," *Management Science* **20**, 1472-1495.

Hart, S. and Z. Levy (1999), "Efficiency Does Not Imply Immediate Agreement," *Econometrica* **67**, 909-912.

Hart, S. and A. Mas-Colell (1989), "Potencial, Value and Consistency," *Econometrica* **57**, 589-614.

Hart, S. and A. Mas-Colell (1996), "Bargaining and Value," *Econometrica* **64**, 357-380.

Herrero, M. (1985), "N-Player Bargaining and Involuntary Underemployment," Chapter IV in *A Strategic Bargaining Approach to Market Institutions*, Ph.D. thesis, London School of Economics.

Herrero, M. (1989), "The Nash Program: Non-Convex Bargaining Problems," *Journal of Economic Theory* **49**, 266-277.

Howard, J. V. (1992), "A Social Choice Rule and its Implementation in Perfect Equilibrium," *Journal of Economic Theory* **56**, 142-159.

Hurwicz, L. (1994), "Economic Design: Adjustment Processes, Mechanisms and Institutions," *Economic Design* **1**, 1-14.

Jackson, M. O. (2001), "A Crash Course in Implementation Theory," *Social Choice and Welfare* **18**, 655-708.

JUN, B. H. (1987), "A Structural Consideration on 3-Person Bargaining," Chapter III in *Essays on Topics in Economic Theory*, Ph. D. Thesis, Department of

Economics, University of Pennsylvania.

Kalai, E. and M. Smorodinsky (1975), “Other Solutions to Nash’s Bargaining Problem,” *Econometrica* **43**, 513-518.

Kalai, E., A. Postlewaite and J. Roberts (1979), “A Group Incentive Compatible Mechanism Yielding Core Allocations,” *Journal of Economic Theory* **20**, 13-22.

Krishna, V. and R. Serrano (1995), “Perfect Equilibria of a Model of n-Person Non-Cooperative Bargaining,” *International Journal of Game Theory* **24**, 259-272.

Krishna, V. and R. Serrano (1996), “Multilateral Bargaining,” *Review of Economic Studies* **63**, 61-80.

Lagunoff, R. (1994), “A Simple Non-Cooperative Core Story,” *Games and Economic Behavior* **7**, 54-61.

LENSBERG, T. (1988), “Stability and the Nash Solution,” *Journal of Economic Theory* **45**, 330-341.

Mas-Colell, A. (1988), “Algunos Comentarios sobre la Teoria Cooperativa de los Juegos,” *Cuadernos Económicos* **40**, 143-161.

Mas-Colell, A. (1989), “An Equivalence Theorem for a Bargaining Set,” *Journal of Mathematical Economics* **18**, 129-139.

Maschler, M. and G. Owen (1989), “The Consistent Shapley Value for Hyperplane Games,” *International Journal of Game Theory* **18**, 389-407.

Maschler, M. and G. Owen (1992), “The Consistent Shapley Value for Games without Side Payments,” in R. Selten (ed.) *Rational Interactions: Essays in Honor of J. C. Harsanyi*, Springer Verlag, Berlin, 5-12.

Maskin, E. (2003), “Bargaining, Coalitions and Externalities,” Mimeo, Institute for Advanced Study, Princeton, Presidential Address to the Econometric Society.

Maskin, E. and T. Sjostrom (2002), “Implementation Theory,” in K. J. Arrow, A. Sen and K. Suzumura (eds.) *Handbook of Social Choice and Welfare (vol. I)*, New York, Elsevier Science B.V.

Miyagawa, E. (2002), “Subgame-Perfect Implementation of Bargaining Solutions,” *Games and Economic Behavior* **41**, 292-308.

Moldovanu, B. and E. Winter (1994), “Increasing Returns to Cooperation and Core Implementation,” *Journal of Mathematical Economics* **23**, 533-548.

Moldovanu, B. and E. Winter (1995), “Order Independent Equilibria,” *Games and Economic Behavior* **9**, 21-34.

Montero, M. (1999), “Coalition Formation in Games with Externalities,” Tilburg CenTER Discussion Paper Series 99-121.

Montero, M. (2002),“Noncooperative Bargaining in Apex Games and the Kernel,” *Games and Economic Behavior* **41**, 309-321.

Moore, J. (1992), “Implementation, Contracts and Renegotiation in Environments with Complete Information," in J. J. Laffont (ed.) *Advances in Economic Theory, VI World Congress of the Econometric Society* (vol. I), Cambridge University Press, Cambridge.

Morelli, M. and M. Montero (2003), "The Demand Bargaining Set: General Characterization and Application to Weighted Majority Games," *Games and Economic Behavior* **42**, 137-155.

Moulin, H. (1984), "Implementing the KalaiSmorodinsky Bargaining Solution," *Journal of Economic Theory* **33**, 32-45.

Muthoo, A. (1999), *Bargaining Theory with Applications*, Cambridge University Press.

Naeve, J. (1999), "Nash Implementation of the Nash Bargaining Solution Using Intuitive Message Spaces," *Economics Letters* **62**, 23-28.

Nash, J. F. (1950), "The Bargaining Problem," *Econometrica* **18**, 155-162.

Nash, J. F. (1953), "Two Person Cooperative Games," *Econometrica* **21**, 128-140.

Okada, A. (1996), "A Noncooperative Coalitional Bargaining Game with Random Proposers," *Games and Economic Behavior* **16**, 97-108.

Osborne, M. J. and A. Rubinstein (1990), *Bargaining and Markets*, Academic Press, San Diego, CA.

Palfrey, T. R. (2002), "Implementation Theory," in R. J. Aumann and S. Hart (eds.) *Handbook of Game Theory with Economic Applications (vol. III)*, New York, Elsevier Science.

Peleg, B. (1986), "On the Reduced Game Property and its Converse," *International Journal of Game Theory* **15**, 187-200.

Peleg, B. and P. Sudholter (2003), *Introduction to the Theory of Cooperative Games*, Kluwer Academic Publishers.

Pérez-Castrillo, D. (1994), "Cooperative Outcomes through Non-Cooperative Games," *Games and Economic Behavior* **7**, 428-440.

Pérez-Castrillo, D. and D. Wettstein (2000), "Implementation of Bargaining Sets via Simple Mechanisms," *Games and Economic Behaviour* **31**, 106-120.

Pérez-Castrillo, D. and D. Wettstein (2001), "Bidding for the Surplus: a Non-Cooperative Approach to the Shapley Value," *Journal of Economic Theory* **100**, 274-294.

Perry, M. and P. Reny (1994), "A Non-Cooperative View of Coalition Formation and the Core," *Econometrica* **62**, 795-817.

Rubinstein, A. (1982), "Perfect Equilibrium in a Bargaining Model," *Econometrica* **50**, 97-109.

Schmeidler, D. (1969), "The Nucleolus of a Characteristic Function Game," *SIAM Journal of Applied Mathematics* **17**, 1163-1170.

Selten, R. (1981), "A Non-Cooperative Model of Characteristic Function Bargaining," in R. Aumann et al. (ed.) *Essays in Game Theory and Mathematical Eco-*

nomics in Honor of Oskar Morgenstern, Bibliographisches Institut Mannheim, 131-151.

Selten, R. and M. Wooders (1991), "A Game Equilibrium Model of Thin Markets," In R. Selten (ed.) *Game Equilibrium Models III; Strategic Bargaining,* Springer-Verlag, Berlin.

Serrano, R. (1993), "Non-Cooperative Implementation of the Nucleolus: the 3-Player Case," *International Journal of Game Theory* **22**, 345-357.

Serrano, R. (1995a), "Strategic Bargaining, Surplus Sharing Problems and the Nucleolus," *Journal of Mathematical Economics* **24**, 319-329.

Serrano, R. (1995b), "A Market to Implement the Core," *Journal of Economic Theory* **67**, 285-294.

Serrano, R. (1997a), "A Comment on the Nash Program and the Theory of Implementation," *Economics Letters* **55**, 203-208.

Serrano, R. (1997b), "Reinterpreting the Kernel," *Journal of Economic Theory* **77**, 58-80.

Serrano, R. (2004), "The Theory of Implementation of Social Choice Rules," *SIAM Review* **46**, 377-414.

Serrano, R. and R. Vohra (1997), "Non-Cooperative Implementation of the Core," *Social Choice and Welfare* **14**, 513-525.

Serrano, R. and R. Vohra (2002a), "Bargaining and Bargaining Sets," *Games and Economic Behavior* **39**, 292-308.

Serrano,R. and R. Vohra (2002b), "Implementing the Mas-Colell Bargaining Set," *Investigaciones Económicas* **26**, (2002), 285-298.

Shapley, L. S. (1953), "A Value for n-Person Games," in *Contributions to the Theory of Games IV,* A.W. Tucker and R.D. Luce (ed.), Princeton, Princeton University Press, 307-317.

Sobolev, A. I. (1975), "The Characterization of Optimality Principles in Cooperative Games by Functional Equations," *Mathematical Methods in the Social Sciences* **6**, 150-165 (in Russian).

Sonmez, T. (1999), "Strategy-Proofness and Essentially Single-Valued Cores," *Econometrica* **67**, 677-689.

Thomson, W. (1990), "The Consistency Principle," in T. Ichiishi, A. Neyman and Y. Tauman (ed.) *Game Theory and Applications*, pp. 187-215, Academic Press, San Diego, CA.

Trockel, W. (1999),"Unique Nash Implementation for a Class of Bargaining Solutions," *International Game Theory Review* **1**, 267-272.

Trockel, W. (2000), "Implementation of the Nash Solution Based on its Walrasian Characterization," *Economic Theory* **16**, 277-294.

Trockel, W. (2002a), "Integrating the Nash Program into Mechanism Theory," *Review of Economic Design* **7**, 27-43.

Trockel, W. (2002b), "Can and Should the Nash Program Be Looked at as a Part of Mechanism Theory?" in Sertel, M. and Koray. S. (ed.) *Advances in Economic Design*, Springer, Heidelberg.

Trockel, W. (2002c), "A Universal Meta Bargaining Realization of the Nash Solution," *Social Choice and Welfare* **19**, 581-586.

Vidal-Puga, J.-J. (2004), "Bargaining with Commitments," Mimeo, Universidad de Vigo, forthcoming in *International Journal of Game Theory*.

von Neumann, J. and O. Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ.

Wilson, R. (1978), "Competitive Exchange," *Econometrica* **46**, 577-585.

Winter, E. (1994a), "The Demand Commitment Bargaining and Snowballing of Cooperation," *Economic Theory* **4**, 255-273.

Winter, E. (1994b), "Non-Cooperative Bargaining in Natural Monopolies," *Journal of Economic Theory* **64**, 202-220.

Winter, E. (1996), "Mechanism Robustness in Multilateral Bargaining," *Theory and Decision* **40**, 131-147.

Yan, H. (2002), "Noncooperative Selection of the Core," *International Journal of Game Theory* **31**, 527-540.

Young, H. P. (1985), "Monotonic Solutions of Cooperative Games," *International Journal of Game Theory* **14**, 65-72.

Young, H. P. (1998), "Cost Allocation, Demand Revelation, and Core Implementation," *Mathematical Social Sciences* **36**, 213-228.