Cooperative Games in Strategic Form*

Sergiu Hart† Andreu Mas-Colell‡

July 28, 2008

Abstract

In this paper we view bargaining and cooperation as an interaction superimposed on a strategic form game. A multistage bargaining procedure for N players, the "proposer commitment" procedure, is presented. It is inspired by Nash's two-player variable-threat model; a key feature is the commitment to "threats." We establish links to classical cooperative game theory solutions, such as the Shapley value in the transferable utility case. However, we show that even in standard pure exchange economies the traditional coalitional function may not be adequate when utilities are not transferable.

1 Introduction

In this paper we take the following point of view on cooperation and bargaining: there is an underlying physical reality represented by a game in strategic form, and bargaining is a noncooperative interaction that acts through an institutional setup superimposed on the strategic form. In a sense the aim

*The authors thank Robert Aumann, Ken Binmore, Jean-François Mertens, Roger Myerson, and Andrew Postlewaite for useful comments and discussions. Research partially supported by grants of: the Israel Academy of Sciences and Humanities; the Fundacion BBVA; the Ministerio de Educación y Ciencia, Spain, through a Consolider Grant; and by the Institute of Advanced Studies at the Hebrew University of Jerusalem.

†Center for the Study of Rationality, Department of Economics, and Institute of Mathematics, The Hebrew University of Jerusalem, Feldman Building, Givat Ram, 91904 Jerusalem, Israel. *E-mail*: `hart@huji.ac.il`; *URL*: `http://www.ma.huji.ac.il/hart`.

‡Department of Economics and Business, Universitat Pompeu Fabra, and Barcelona Graduate School of Economics, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain. *E-mail*: `andreu.mas-colell@upf.edu`.

of the institutions is to induce bargaining processes that lead to a selection of final play that is efficient.

Our outlook is inspired both by cooperative game theory and by general equilibrium theory. As in cooperative game theory we aim to handle situations with an arbitrary number of players and recognize that allowing for the possibility of partial breakdown of the negotiation is of the essence (in cooperative games, partial breakdown is captured by the specification of what coalitions other than the grand coalition can obtain). From general equilibrium theory we import the paradigm of a sharp distinction between the underlying data (preferences, endowments, and technologies in general equilibrium, the strategic form here) and the institutions that operate in them (typically markets in general equilibrium, bargaining procedures here).

We are not the first to argue that a general theory of bargaining should be built over a strategic form. It was already the position taken by Nash (1953) in his proposal for the endogeneization of the threat points of his axiomatic theory of bargaining (Nash, 1950), a proposal that was extended to N-player situations by Harsanyi (1959, 1963). As for cooperative game theory, the need of a strategic-form foundation has been persistently felt. One suggestion was provided for the transferable-utility ("TU") case by the founders (von Neumann and Morgenstern, 1944, proposed to define what a coalition could reach as the maximin level for the sum of the payoffs of the members of the coalition), and later generalized (Aumann, 1959) to the non-transferable-utility ("NTU") case in the guise of the "alpha" and "beta" coalitional forms. Dissatisfaction with these definitions drew attention to particular classes of games where the determination of the coalitional form appeared uncontroversial (the "c-games" of Shapley–Shubik, a leading example of which is exchange economies; see Shubik, 1983, Section 6.2.2), and thus the theory of bargaining could be nicely factored through the coalitional form. We will have some opportunity to debate this point (see Section 7 below).

We shall make a specific proposal for a bargaining procedure, what we call the "proposer commitment" (PC) procedure. We do not pretend that it is the most general procedure since plainly it is not, but we do believe that it is comparatively simple and that in applications it captures some important features.

The PC procedure is inspired by Nash (1953) in a crucial feature: we assume that players can commit to threats. We view this commitment possibility as going hand in hand with the set of strong institutions that must be in place if, as we assume, the outcome of bargaining is enforceable. But we depart from Nash (1953) in an important aspect: at each step of the negotiations only one player (the "proposer") makes threats. We do so because we want a bargaining procedure that, in the spirit of modern bargaining theory (see, e.g., Binmore, Osborne and Rubinstein, 1992), has as players the players of the original, underlying strategic form. In the two-player setting of Nash (1953), his simultaneous threats model can be made to pass this test (replace, for example, the axiomatically based part of his solution by bargaining in the style of Stahl–Rubinstein; see, e.g., Osborne and Rubinstein, 1990, and Houba and Bolt, 2002), but the N-player generalization of Harsanyi (1959, 1963) does not, at least to our knowledge (Harsanyi defines a sort of noncooperative bargaining, but it is between fictitious players, one for each coalition). Thus in a sense (reminiscent of Shapley, 1969) we could present our solution as a simplification of the Nash–Harsanyi approach.

There is another source of inspiration for the PC approach: the bargaining procedure formulated by ourselves (Hart and Mas-Colell, 1996a) for the context where the underlying reality is a game in coalitional form. The sequential nature of the announcements and proposals we take from there, along with the idea that a rejected proposer becomes passive for the rest of the game (with some probability). But the consideration of a strategic form as the underlying reality allows us now to enrich the determination of what happens with the play of the rejected proposer.

The paper is organized as follows. In Section 2 the basic model and the PC bargaining procedure are presented. In Section 3 we establish the standard existence and optimality properties. In Section 4 we focus on two special cases: two-person games and games with transferable utility. In Section 5 we discuss, in a particular context where the threats turn out to be "fixed" in equilibrium, a general connection of the PC solution with the Shapley value in the TU case. In Sections 6 and 7 we reexamine the concept of c-games (in particular the standard class of exchange economies) as candidates for the sort of simplification that would allow the factorization of bargaining analysis through a coalitional form (deducible from the fundamentals of the game). We discover an important difference between TU games—for which the simplification is possible (see Section 6)—and general NTU games—for which it is not. We show the latter in Section 7 by means of an example related to the transfer paradoxes of general equilibrium theory.

2 The Model

The basic data is an *N-person game in strategic form* $G = (N, (A^i)_{i\in N}, (u^i)_{i\in N})$, where N is a finite set of players, and each player $i \in N$ has a finite set of actions A^i and a payoff function $u^i : A \to \mathbb{R}$, with $A := \prod_{i\in N} A^i$. A *mixed action* of player i is $x^i \in \Delta(A^i)$, where $\Delta(A^i) = \{(x^i(a^i) \in \mathbb{R}_+^{A^i} : \sum_{a^i\in A^i} x^i(a^i) = 1\}$ is the probability simplex on A^i.

For each set of players $S \subset N$ (a *coalition*), let $A^S := \prod_{i\in S} A^i$ denote the set of pure action combinations of the members of S. A *correlated action* of S is $z^S \in \Delta(A^S)$, a probability distribution on pure action combinations of S. The payoff functions are as usual multilinearly extended to mixed and correlated actions.

2.1 The Proposer-Commitment (PC) Procedure

We now introduce the basic bargaining procedure.

Let $0 \leq \rho < 1$ be a fixed parameter; think of ρ as the probability of "repeat." The bargaining proceeds in rounds. In each round there is a set $S \subset N$ of "active" players, the actions of each "inactive" player $j \notin S$ being fixed at some $b^j \in A^j$; put $b^{N\setminus S} = (b^j)_{j\in N\setminus S}$. We will refer to $\omega = (S, b^{N\setminus S})$ as a *state*. Initially, everyone is active, i.e., $S = N$ (and so the starting state is $(N, \cdot)$). Each round, with state $\omega = (S, b^{N\setminus S})$, proceeds as follows.

1. A "proposer" $k \in S$ is selected out of S at random, with all members of S being equally likely to be selected.

2. The proposer k chooses a pair (z^S, x^k), where $z^S \in \Delta(A^S)$ is a correlated action of S and $x^k \in \Delta(A^k)$ is a mixed action of player k; think of z^S as a "proposed agreement" for S, and of x^k as a "threat."

3. Each player in S is asked, in some order (deterministic or random), whether he accepts or rejects the proposed agreement z^S.

4. If they all agree to z^S, then the procedure ends as follows: a joint action $a^S \in A^S$ is selected according to the distribution z^S, and the N-tuple of actions $(a^S, b^{N\backslash S}) \in A$ is played in the original strategic game G.

5. If at least one player in S rejects z^S, then with probability ρ the state does not change (it remains ω=$(S, b^{N\backslash S})$; we call this "repeat"), and with probability $1 - \rho$ the rejected proposer k becomes inactive.

6. If the rejected proposer becomes inactive, then the randomization x^k is performed; let $b^k \in A^k$ be its realization. The action of player k is fixed from now on at $b^k \in A^k$, and the new state is $\omega' = (S\backslash\{k\}, (b^{N\backslash S}, b^k))$: the set of active players is $S\backslash\{k\}$ and the actions of the inactive players are $(b^{N\backslash S}, b^k)$.

7. A new round is started (i.e., one goes back to step 1), with the state being the same ω in case of repeat, and ω' as in step 6 otherwise.

2.2 Outcomes and Equilibria

We are interested in the (*subgame-*)*perfect equilibria* of the PC procedure that are, in addition, as simple as possible, i.e., *stationary*. This means that the decisions of the players depend only on the payoff-relevant variables, not on the history nor on the calendar time. Formally, for each state $\omega = (S, b^{N\backslash S})$ and proposer $k \in S$, the announcement (z^S, x^k) of player k depends only on S, $b^{N\backslash S}$, and k, and the decision of each player $i \in S\backslash k$ to accept or reject depends only on S, $b^{N\backslash S}$, k, z^S, x^k, and i. Stationary subgame-perfect equilibria will be called *SP equilibria* for short.

For simplicity, we will assume that each player uses the tie-breaking rule of accepting a proposal when accepting it and rejecting it give him the same expected payoff.

The play of the PC procedure ends with probability one (since $\rho < 1$); its end result is an N-tuple of actions $a \in A$ in the original game G (see step 4 in the PC procedure), which we call the *final* N-tuple of actions. This final a is random: it depends on the randomizations of nature (e.g., selecting the proposers and repeating or not after rejection) and of the players themselves.

Fix an N-tuple of stationary strategies $\sigma = (\sigma^i)_{i \in N}$.

For each state $\omega = (S, b^{N \backslash S})$, let $\alpha_\omega \in \Delta(A)$ denote the probability distribution of the final N-tuple of actions in the subgame starting from state ω. Since the actions of the players outside S are fixed at $b^{N \backslash S}$, the randomness affects only the actions of the players in S, and so $\alpha_\omega = \zeta^S_\omega \times \{b^{N \backslash S}\}$ for some $\zeta^S_\omega \in \Delta(A^S)$. We refer to ζ^S_ω as the *outcome of state* ω. Similarly, $\zeta^S_{\omega,k} \in \Delta(A^S)$ denotes the probability distribution of the final actions of S after $k \in S$ has been selected as proposer; we call it the *outcome of state* $\omega = (S, b^{N \backslash S})$ *and proposer* k. Since the proposer is equally likely to be any member of S, we have[1]

$$\zeta^S_\omega = \frac{1}{|S|} \sum_{k \in S} \zeta^S_{\omega,k} \tag{1}$$

for every state $\omega = (S, b^{N \backslash S})$. The collection of outcomes ζ^S_ω and $\zeta^S_{\omega,k}$ for all possible states and proposers (i.e., $\omega = (S, b^{N \backslash S})$ for $S \subset N$, $b^{N \backslash S} \in A^{N \backslash S}$, and $k \in S$) will be called an *outcome configuration* (obtained from σ) and will be denoted $\boldsymbol{\zeta}$.

For every $k \in S$ and every $b^k \in A^k$, let $(\omega || b^k) := (S \backslash k, (b^{N \backslash S}, b^k))$ denote the state obtained from ω when k becomes inactive and his action is fixed at b^k; for every $x^k \in \Delta(A^k)$ let

$$\eta^S_{\omega,k}(x^k) := \sum_{b^k \in A^k} x^k(b^k) \left(\zeta^{S \backslash k}_{(\omega || b^k)} \times \{b^k\} \right) \in \Delta(A^S) \tag{2}$$

be the expected outcome for S following the implementation of the threat

[1] For a finite set Z, we denote by $|Z|$ the number of elements of Z.

x^k. We will say that an announcement $(z^S, x^k) \in \Delta(A^S) \times \Delta(A^k)$ of player k is "acceptable" if, when the continuation is according to $\boldsymbol{\zeta}$, each responder's payoff from accepting z^S is no less than his payoff from rejecting it, i.e.,

$$u^i(z^S, b^{N\backslash S}) \geq \rho u^i(\zeta^S_\omega, b^{N\backslash S}) + (1-\rho)\, u^i(\eta^S_{\omega,k}(x^k), b^{N\backslash S}) \tag{3}$$

for every $i \in S\backslash k$ (recall that after rejection with probability ρ the state remains ω and with probability $1-\rho$ player k becomes inactive and his threat is implemented). Let $Y \equiv Y_{\omega,k}(\boldsymbol{\zeta})$ denote the set of acceptable announcements of k :

$$Y := \{(z^S, x^k) \in \Delta(A^S) \times \Delta(A^k) : (3) \text{ holds for every } i \in S\backslash k\},$$

and let $Y^* \equiv Y^*_{\omega,k}(\boldsymbol{\zeta})$ be the set of those acceptable announcements that maximize the payoff of the proposer[2] k:

$$Y^* := \arg \max_{(z^S, x^k) \in Y} u^k(z^S, b^{N\backslash S}).$$

Finally, denote by $Z \equiv Z_{\omega,k}(\boldsymbol{\zeta})$ and $Z^* \equiv Z^*_{\omega,k}(\boldsymbol{\zeta})$ the projections of the sets Y and Y^*, respectively, on the z^S-coordinate:

$$\begin{aligned} Z &:= \{z^S \in \Delta(A^S) : (z^S, x^k) \in Y \text{ for some } x^k \in \Delta(A^k)\}; \\ Z^* &:= \{z^S \in \Delta(A^S) : (z^S, x^k) \in Y^* \text{ for some } x^k \in \Delta(A^k)\}. \end{aligned}$$

We claim that the SP equilibrium conditions on the outcome configuration $\boldsymbol{\zeta}$ can be stated as

$$\zeta^S_{\omega,k} \in Z^*_{\omega,k}(\boldsymbol{\zeta}) \tag{4}$$

for every state $\omega = (S, b^{N\backslash S})$ and $k \in S$. Note that (4) is a fixed-point-type condition.

Proposition 1 *An outcome configuration* $\boldsymbol{\zeta}$ *is obtained from an SP equilibrium of the PC procedure if and only if* $\boldsymbol{\zeta}$ *satisfies condition (4) for all states* $\omega = (S, b^{N\backslash S})$ *and* $k \in S$.

[2] We write arg max for the set of maximizers.

Proof. Let $\boldsymbol{\zeta}$ satisfy the conditions (4). Define an N-tuple of strategies σ as follows: in state $\omega = (S, b^{N\backslash S})$, when $k \in S$ is the proposer he announces an element $(\tilde{z}^S, \tilde{x}^k) \in Y^*_{\omega,k}$ with $\tilde{z}^S = \zeta^S_{\omega,k}$, and when $i \in S$ is a responder he accepts a proposal z^S if and only if (3) holds. It is straightforward to verify (using a "one-deviation property" as in, e.g., Osborne and Rubinstein, 1994, Lemma 98.2) that σ constitutes an SP equilibrium, and its outcome configuration is precisely the given $\boldsymbol{\zeta}$.

Conversely, let σ be an SP equilibrium with outcome configuration $\boldsymbol{\zeta}$. Take a state $\omega = (S, b^{N\backslash S})$ and a proposer $k \in S$, and consider a single deviation from σ, at this point only, by player k. We claim that the set of outcomes that k can induce is precisely $Z \equiv Z_{\omega,k}(\boldsymbol{\zeta})$. Indeed, an announcement (z^S, x^k) yields acceptance by all players in $S\backslash k$ if and only if conditions (3) hold, i.e., if and only if $(z^S, x^k) \in Y$ (by the equilibrium requirement when there is strict inequality, and by the tie-breaking rule when there is equality). When (z^S, x^k) is rejected the continuation outcome is $\bar{z}^S := \rho\zeta^S_\omega + (1-\rho)\,\eta^S_{\omega,k}(x^k)$, and here too we have $(\bar{z}^S, x^k) \in Y$ (conditions (3) hold as equalities). Therefore we have shown that Z, the projection of Y on the z^S-coordinate, is indeed the set of all possible outcomes that k can induce at this point (whether there is acceptance or rejection). But k maximizes his payoff (since σ is an equilibrium), from which condition (4) follows. $\square$

We note two simple but useful facts. For every state $\omega = (S, b^{N\backslash S})$ let[3]

$$C(\omega) := \{u^S(z^S, b^{N\backslash S}) \in \mathbb{R}^S : z^S \in \Delta(A^S)\}$$

be the set of feasible payoff vectors for the coalition S at ω (i.e., given the fixed actions $b^{N\backslash S} \in A^{N\backslash S}$ of the players outside S).

Lemma 2 *Let σ be an SP equilibrium with outcome configuration $\boldsymbol{\zeta}$. For every state $\omega = (S, b^{N\backslash S})$ and $k \in S$:*

(i) *$Y_{\omega,k}(\boldsymbol{\zeta})$ is a nonempty polytope; and*

[3] We write $u^S(z)$ for the payoff vector $(u^i(z))_{i\in S} \in \mathbb{R}^S$.

(ii) *there does not exist* $c \in C(\omega)$ *such that* $c \geq u^S(\zeta^S_{\omega,k})$ *with strict inequality* $c^k > u^k(\zeta^S_{\omega,k})$ *in the k-th coordinate.*[4]

Proof. (i) The set Y is nonempty since for every $x^k \in \Delta(A^k)$ we have $(\bar{z}^S, x^k) \in Y$, where $\bar{z}^S := \rho\zeta^S_\omega + (1-\rho)\,\eta^S_{\omega,k}(x^k)$ (conditions (3) hold as equalities). It is a convex polytope since it is defined by the finitely many inequalities (3) that are linear in z^S and x^k (the outcomes ζ^S and $\zeta^{S\setminus k}$ are fixed).

(ii) Assume there is $z^S \in \Delta(A^S)$ such that $c = u^S(z^S, b^{N\setminus S}) \in C(\omega)$ satisfies $c \geq u^S(\zeta^S_{\omega,k})$ and $c^k > u^k(\zeta^S_{\omega,k})$. Replacing $\zeta^S_{\omega,k}$ by z^S preserves the inequalities (3): indeed, the left-hand side increases by $\delta := c^i - u^i(\zeta^S_{\omega,k}) \geq 0$, whereas the right-hand side increases by less than δ, specifically $(\rho/|S|)\delta$; see (1). Therefore $(z^S, x^k) \in Y$ is also an acceptable announcement (with the threat x^k unchanged), but the payoff of k is strictly higher there, which contradicts (4). $\square$

3 General Results

In this section we prove two general results of a standard type. First, we show that SP equilibria exist; and second, that as the probability of repeat gets close to 1—that is, as the "cost of delay" goes to zero—the SP equilibrium outcomes approach Pareto efficiency.

3.1 Existence

Proposition 3 *There exists an SP equilibrium.*

Proof. We proceed by induction on S. For $|S| = 1$, say $S = \{i\}$, the strategy of player i in state $(\{i\}, b^{N\setminus i})$ consists of choosing $z^i \in \arg\max_{x^i \in \Delta(A^i)} u^i(x^i, b^{N\setminus i})$.

Let the state be $\omega = (S, b^{N\setminus S})$, and assume that equilibrium strategies and outcomes have been determined for all states $\omega' = (S', b^{N\setminus S'})$ with $S' \subsetneq S$.

[4] This implies that $u^S(\zeta^S_{\omega,k})$ is *weakly* Pareto efficient in $C(\omega)$.

For each $c \in C(\omega)$ (the set of feasible payoff vectors for S) and $k \in K$, let

$$\begin{aligned}\Phi_k(c) := \{&(z^S, x^k) \in \Delta(A^S) \times \Delta(A^k) : \\ &u^i(z^S, b^{N\backslash S}) - (1-\rho)u^i(\eta^S_{\omega,k}(x^k)) \geq \rho c^i \text{ for all } i \in S\backslash k\},\end{aligned}$$

where $\eta^S_{\omega,k}(x^k)$ is defined in (2) (based on the $\zeta^{S\backslash k}_{(\omega||b^k)}$, which have already been determined by induction). The set $\Phi_k(c)$ is nonempty (take $z^S \in \Delta(A^S)$ with $u^S(z^S, b^{N\backslash S}) = c$; then $(\tilde{z}^S, x^k) \in \Phi_k(c)$ where $\tilde{z}^S = \rho z^S + (1-\rho)\eta^S_{\omega,k}(x^k)$) and is a convex polytope (note that $\eta^S_{\omega,k}(x^k)$ is linear in x^k); the correspondence Φ_k is continuous on $C(\omega)$ (by Lemma 4 below). Therefore

$$\Phi^*_k(c) := \arg \max_{(z^S,x^k)\in\Phi_k(c)} u^k(z^S, b^{N\backslash S})$$

is a nonempty, convex-valued, and upper-semicontinuous correspondence (the latter by the Maximum Theorem since u^k is linear and thus continuous, and Φ_k is a continuous correspondence; see, e.g. Hildenbrand 1974, Corollary to Theorem B.III.4). Hence the same holds for the correspondences Ψ_k and Ψ, defined by

$$\begin{aligned}&\Psi_k(c) := \{u^S(z^S, b^{N\backslash S}) : (z^S, x^k) \in \Phi^*_k(c)\} \text{ and} \\ &\Psi(c) := \frac{1}{|S|}\sum_{k\in S} \Psi_k(c).\end{aligned}$$

We can therefore apply Kakutani's Fixed-point Theorem (see, e.g., Hildenbrand 1974, C.III (14)) to the correspondence Ψ (with domain $C(\omega)$), to obtain $\bar{c} \in C(\omega)$ with $\bar{c} \in \Psi(\bar{c})$. This yields, in turn, $\bar{c}_k \in \Psi_k(\bar{c})$ with $\bar{c} = (1/|S|)\sum_{k\in S} \bar{c}^k$, and $(\bar{z}^S, \bar{x}^k) \in \Phi^*_k(\bar{c})$. It is immediate to verify that the announcements $(\bar{z}^S, \bar{x}^k)$ for all $k \in S$ constitute equilibrium announcements in state ω. This completes the induction step, and thus proves our claim. $\square$

Remark. When $\rho = 0$ there is no need to use a fixed-point theorem to prove existence: the SP equilibria can be computed recursively, starting with singleton S.

In the proof we have used the following:

Lemma 4 *Let D be an $m \times n$ matrix and put $F(w) := \{x \in \mathbb{R}^n : Dx \geq w\}$ for every $w \in \mathbb{R}^m$. Then F is a continuous correspondence on $W := \{w \in \mathbb{R}^m : F(w) \neq \emptyset\}$.*

Proof. Upper-semicontinuity is immediate. For lower-semicontinuity, let $x_0 \in \mathbb{R}^n$ satisfy $Dx_0 \geq w_0$, and let $w_r \to w_0$ with $w_r \in W$ for all r; we have to show that for every r there is x_r with $Dx_r \geq w_r$. It suffices to consider the case where only one coordinate of w_0 changes, say, $w_r = w_0 + (\delta_r, 0, ..., 0)$. If $\delta_r \to 0^-$, then take $x_r = x_0$. If $\delta_r \to 0^+$, then let x_1 satisfy $Dx_1 \geq w_1$ (recall that $w_1 \in W$), and then $x_r := (1 - \delta_r/\delta_1)x_0 + (\delta_r/\delta_1)x_1$ satisfies $Dx_r \geq w_0 + (\delta_r, 0, ..., 0) = w_r$. □

3.2 Pareto Efficiency

In equilibrium, every individual proposal $\zeta^S_{\omega,k}$ is (weakly) Pareto efficient (see Lemma 2 (ii)). Therefore the outcomes ζ^S_ω may fail to be efficient only if the Pareto-efficient boundary is not a hyperplane and the individual proposals of different proposers are different (see (1)). However, if ρ is close to 1—i.e., the "cost of delay" is small—then the early-proposer's advantage will be small, and so the individual proposals will be similar and their average almost Pareto efficient.

To see this, let $\boldsymbol{\zeta}(\rho)$ be an SP equilibrium outcome for the PC bargaining procedure with parameter ρ—we will refer to it as the PC "ρ-procedure." Consider a limit point $\bar{\boldsymbol{\zeta}}$ of $\boldsymbol{\zeta}(\rho)$ as $\rho \to 1$ (i.e., there is a sequence $\rho_m \to 1$ such that $\boldsymbol{\zeta}(\rho_m) \to \bar{\boldsymbol{\zeta}}$ as $m \to \infty$). Then:

Theorem 5 *Let $\bar{\boldsymbol{\zeta}} = (\bar{\zeta}^S_\omega)_\omega$ be a limit point as $\rho \to 1$ of SP equilibrium outcomes $\boldsymbol{\zeta}(\rho) = (\zeta^S_\omega(\rho))_\omega$ of the PC ρ-procedures. Then for every state $\omega = (S, b^{N\backslash S})$ the limit outcome $\bar{\zeta}^S_\omega$ in state ω is Pareto efficient for S given $b^{N\backslash S}$.*

Proof. Assume for simplicity that $\boldsymbol{\zeta}(\rho) \to \bar{\boldsymbol{\zeta}}$ as $\rho \to 1$ (otherwise restrict the arguments to the sequence ρ_m with $\boldsymbol{\zeta}(\rho_m) \to \bar{\boldsymbol{\zeta}}$). Put $g_{\omega,k} \equiv g_{\omega,k}(\rho) := u^S(\zeta^S_{\omega,k}(\rho), b^{N\backslash S})$, $g_\omega \equiv g_\omega(\rho) := u^S(\zeta^S_\omega(\rho), b^{N\backslash S})$, and $\bar{g}_\omega := u^S(\bar{\zeta}^S_\omega, b^{N\backslash S})$ for all $k \in S$; thus $g_\omega \to \bar{g}_\omega$ as $\rho \to 1$. Let M be a bound on all possible payoffs of all players; for each ρ we have

$$g^i_{\omega,k} \geq \rho g^i_\omega + (1-\rho)u^i(\eta^S_{\omega,k}(x^k), b^{N\setminus S}) \geq g^i_\omega - (1-\rho)2M \tag{5}$$

(for $i \neq k$ it follows from (4), the definition of Y, and (3); for $i = k$, from (4) together with $(\tilde{z}^S, x^k) \in Y$ for $\tilde{z}^S := \rho\zeta^S_\omega + (1-\rho)\,\eta^S_{\omega,k}(x^k))$. Now $g^i_\omega = (1/|S|)\sum_{k\in S} g^i_{\omega,k}$ by (1), and so adding the inequalities (5) for all k except some $k_0 \in S$ (keep i fixed) yields

$$|S|g^i_\omega - g^i_{\omega,k_0} \geq (|S|-1)(g^i_\omega - (1-\rho)2M),$$

or

$$g^i_{\omega,k_0} \leq g^i_\omega + (1-\rho)(|S|-1)2M \leq g^i_\omega + (1-\rho)2M|N|$$

for all $k_0 \in S$. Thus

$$-(1-\rho)2M \leq g^i_{\omega,k} - g^i_\omega \leq (1-\rho)2M|N|$$

(replace k_0 by k to get the second inequality, and recall (5) for the first); hence, as $\rho \to 1$, we get $g^i_{\omega,k} - g^i_\omega \to 0$, which, since $g^i_\omega \to \bar{g}^i_\omega$, implies that $g^i_{\omega,k} \to \bar{g}^i_\omega$ for all $i, k \in S$.

If $\bar{g}_\omega$ is not Pareto efficient in $C(\omega)$, then there exist $k \in S$ and $c \in C(\omega)$ such that $c \geq \bar{g}_\omega$, with strict inequality in the k-th coordinate. Then $c(\rho) := g_{\omega,k}(\rho) + (1/2)(c - \bar{g}_\omega)$ satisfies $c(\rho) \geq g_{\omega,k}(\rho)$, with strict inequality in the k-th coordinate; also, for ρ close enough to 1, we have $c(\rho) \in C$ (use the fact that C is determined by finitely many linear inequalities and $g_{\omega,k}(\rho) \to \bar{g}_\omega$). But this contradicts Lemma 2 (ii). $\square$

4 Two Reference Cases: Two Players and Transferable Utility

In this section we spell out the nature of our solution for two simple and classical cases.

4.1 Two-Person Games

The SP equilibria of the PC procedure relate very directly to the Nash bargaining solution for the case of two players (cf. Hart and Mas-Colell 1996a; see also Houba and Bolt, 2002, for more on two-person bargaining games).

Given a two-person game G with $N = \{1, 2\}$, for each player $i \in N$ let q^i be the payoff level that the other player $j \neq i$ can hold i to, by using pure strategies; i.e.,

$$q^i := \min_{a^j \in A^j} \max_{a^i \in A^i} u^i(a^i, a^j).$$

Let

$$D := \{u^N(z) : z \in \Delta(A)\} = \operatorname{co}\{(u^1(a), u^2(a)) : a \in A\}$$

be the set of feasible payoff vectors. (D, q) is called a two-person *pure bargaining problem*, where D is the set of "feasible agreements" and q the "disagreement point," if $q \in D$ and there exists $d \in D$ such that $d^1 > q^1$ and $d^2 > q^2$ (see Nash 1950).

Proposition 6 *Let G be a two-person strategic game such that (D, q) is a pure bargaining problem. If $\bar{\zeta}^N$ is a limit point as $\rho \to 1$ of SP equilibrium outcomes, then $u^N(\bar{\zeta}^N)$ is the Nash bargaining solution of (D, q).*

Proof. $\bar{\zeta}^N$ is Pareto efficient by Theorem 5. If the Pareto boundary $\partial_+ D$ of D contains only $u^N(\bar{\zeta}^N)$ then we are done, since the Nash solution is Pareto efficient. If not, assume first that $u^N(\bar{\zeta}^N)$ is an interior point of $\partial_+ D$. Let $\omega = (N, \cdot)$ be the starting state; since $u^N(\zeta^N_{\omega,1}(\rho)$ is weakly Pareto efficient and it converges to $u^N(\bar{\zeta}^N)$, it follows that $u^N(\zeta^N_{\omega,1}(\rho)) \in \partial_+ D$ for all ρ close enough to 1. For every $b^1 \in A^1$, the payoff that player 2 gets in the state $(\{2\}, b^1)$ is $v^2(b^1) := \max_{a^2 \in A^2} u^2(a^2, b^1)$, and so condition (4) says that $u^1(\zeta^N_{\omega,1}(\rho))$ maximizes $u^1(z^N)$ subject to $u^2(z^N) \geq \rho u^2(\zeta^N(\rho)) + (1-\rho) v^2(b^1)$. Therefore any $b^1 \in A^1$ that has positive probability in the threat $x^1 \in \Delta(A^1)$ that is used by player 1 must make $v^2(b^1)$ as small as possible (here we use the Pareto efficiency of $u^N(\zeta^N_{\omega,1}(\rho)) \in \partial_+ D$); but $\min_{b^1 \in A^1} v^2(b^1) = q^2$, and so $u^1(\zeta^N_{\omega,1}(\rho))$ maximizes $u^1(z^N)$ subject to $u^2(z^N) = \rho u^2(\zeta^N(\rho)) + (1-\rho) q^2$. A similar argument applies when we interchange the two players; from this it follows by standard arguments that the limit $u^N(\bar{\zeta}^N)$ of $u^N(\zeta^N(\rho))$ as $\rho \to 1$ is precisely the Nash bargaining solution of (D, q); see for example Hart and Mas-Colell (1996a, Theorem 3). Finally, if $u^N(\bar{\zeta}^N)$ is an extreme point of $\partial_+ D$—it minimizes, say, player 1's payoff and maximizes player 2's payoff on $\partial_+ D$—then the above argument applies only to $\zeta^N_{\omega,1}(\rho)$; for $\zeta^N_{\omega,2}(\rho)$ we get some $\hat{q}^1(\rho) \geq q^1$ such that $u^2(\zeta^N_{\omega,2}(\rho))$ maximizes $u^2(z^N)$ subject to $u^1(z^N) = \rho u^1(\zeta^N(\rho)) + (1-\rho)\hat{q}^1(\rho)$. Therefore $u^N(\bar{\zeta}^N)$ is the Nash bargaining solution of $(D, (\hat{q}^1, q^2))$ where $\hat{q}^1 \geq q^1$ is a limit point of $\hat{q}^1(\rho)$ as $\rho \to 1$; given that $u^N(\bar{\zeta}^N)$ is that extreme point of $\partial_+ D$ where player 1's payoff is minimal, it easily follows that $u^N(\bar{\zeta}^N)$ is also the solution of (D, q). □

Remark. One could well have $q \notin D$; for example, in the "matching-pennies" game, D is the line segment connecting $(1, -1)$ and $(-1, 1)$, and $q = (1, 1)$. In this case we have a "reverse pure bargaining problem" and $u^N(\bar{\zeta}^N)$ is its solution (see the discussion in Section 4 of Hart and Mas-Colell 1996a; $u^N(\bar{\zeta}^N)$ in this example is $(0, 0)$).

4.2 Transferable Utility

Given the game G, the *individual rational level* in pure actions of player i is

$$r^i := \max_{a^i \in A^i} \min_{a^{N\backslash i} \in A^{N\backslash i}} u^i(a^i, a^{N\backslash i}) = \max_{a^i \in A^i} \min_{z^{N\backslash i} \in \Delta(A^{N\backslash i})} u^i(a^i, z^{N\backslash i});$$

this is the maximum that i can guarantee by using pure strategies. The payoff of player i in any equilibrium of the PC procedure will always be at least r^i (the following strategy σ^i guarantees r^i: when i is the proposer his threat is some $a^i \in A^i$ where the max is attained (i.e., $\min_{a^{N\backslash i} \in A^{N\backslash i}} u^i(a^i, a^{N\backslash i}) = r^i$ holds), and when he is the responder he never accepts any payoff less than[5] r^i). Thus, only payoff vectors $c = (c^i)_{i \in N}$ that are individually rational—i.e., $c^i \geq r^i$ for each i—matter.

We say that the game G is a *strategic game with transferable utility* (a

[5]The intuitive reason why r^i is based on *pure* actions of i (rather than mixed ones) is that if i's proposal is rejected then the randomization in his threat $x^i \in \Delta(A^i)$ is realized, and from then on i is fixed at a pure action that is known to everyone.

"strategic TU game" for short) if for every state $\omega = (S, b^{N\backslash S})$, i.e., for every $S \subset N$ and $b^{N\backslash S} \in A^{N\backslash S}$, there exists a number $v(\omega) \equiv v(S, b^{N\backslash S})$ such that every Pareto efficient and individually rational payoff vector c in $C(\omega) := \{u^S(z^S, b^{N\backslash S}) : z^S \in \Delta(A^S)\}$ satisfies

$$\sum_{i \in S} c^i = v(\omega) \equiv v(S, b^{N\backslash S}). \tag{6}$$

If G is a strategic TU game, then the SP equilibria of the PC procedure become relatively simple to determine. In particular, no fixed point is needed and the computation is not recursive, as the threats can be determined independently for each coalition S.

For every state $\omega = (S, b^{N\backslash S})$, proposer $k \in S$, and mixed action $x^k \in \Delta(A^k)$, extend the definition of $v(\cdot)$ to mixed actions:

$$v(S\backslash k, (b^{N\backslash S}, x^k)) := \sum_{b^k \in A^k} x^k(b^k)\, v(S\backslash k, (b^{N\backslash S}, b^k)),$$

and define

$$\tau_{\omega,k} := \min_{x^k \in A^k} v(S\backslash k, (b^{N\backslash S}, x^k)) = \min_{b^k \in A^k} v(S\backslash k, (b^{N\backslash S}, b^k)); \text{ and} \tag{7}$$

$$X^k_\omega := \arg\min_{x^k \in A^k} v(S\backslash k, (b^{N\backslash S}, x^k)) \tag{8}$$

(note that X^k_ω consists of all pure actions $b^k \in A^k$ that are minimizers of $v(S\backslash k, (b^{N\backslash S}, b^k))$, along with all their probabilistic mixtures). Finally, put

$$D^k v(\omega) := v(\omega) - \tau_{\omega,k} = v(S, b^{N\backslash S}) - \min_{b^k \in A^k} v(S\backslash k, (b^{N\backslash S}, b^k)); \tag{9}$$

this is the "marginal contribution" of player $k \in S$ in state $\omega = (S, b^{N\backslash S})$.

A *threat configuration* $\mathbf{x} = (x^k_\omega)_{\omega,k}$ is a collection of mixed actions $x^k_\omega \in \Delta(A^k)$ for every state $\omega = (S, b^{N\backslash S})$ and every $k \in S$; every N-tuple of stationary pure strategies σ generates such an $\mathbf{x}$: take the second coordinate of the announcements (in state ω a proposer k announces $(\zeta^S_{\omega,k}, x^k_\omega) \in \Delta(A^S) \times \Delta(A^k)$).

Next, let Π be the set of all $|N|!$ orders of the players. For each order $\pi =$

$(i_1, i_2, ..., i_N) \in \Pi$ and threat configuration $\mathbf{x}$ define a probability distribution $Q_\pi \equiv Q_\pi^{\mathbf{x}}$ on A as follows:

$$Q_\pi(b) := \prod_{m=1}^{N} x_{\omega_m}^{i_m}(b^{i_m}) \tag{10}$$

for every $b \in A$, where, for each $m = 1, 2, ..., N$, we put $S_m := \{i_m, i_{m+1}, ..., i_N\}$, $b^{N \setminus S_m} := (b^{i_1}, b^{i_2}, ..., b^{i_{m-1}})$, and $\omega_m := (S_m, b^{N \setminus S_m})$. Taking the order $\pi \in \Pi$ to be random, with all $|N|!$ orders equally likely, yields a joint probability distribution $Q \equiv Q^{\mathbf{x}}$ on $\Pi \times A$:

$$Q(\pi, b) := \frac{1}{N!} Q_\pi(b) \tag{11}$$

for every $\pi \in \Pi$ and $b \in A$. For each $(\pi, b) \in \Pi \times A$ and player $i \in N$, let P_π^i denote the set of predecessors of i in the order π, and let $\omega_{\pi,b}^i = (N \setminus P_\pi^i, b^{P_\pi^i})$ be the state where each predecessor $j \in P_\pi^i$ has his action fixed at the corresponding b^j.

Finally, let

$$\phi^i := \mathbf{E}\left[D^i v(\omega_{\pi,b}^i)\right] \tag{12}$$

be the "expected marginal contribution" of player i to his predecessors, where $\mathbf{E}$ denotes expectation with respect to the distribution $Q^{\mathbf{x}}$ on $\Pi \times A$, and $\omega_{\pi,b}^i$ is the state determined as above.

Proposition 7 *Let G be a strategic TU game with associated function v. If σ is an SP equilibrium of the PC ρ-procedure, then the resulting threat configuration $\mathbf{x} = (x_\omega^k)_{\omega,k}$ satisfies $x_\omega^k \in X_{\omega,k}$ for every ω and k (see (7) and (8)). Conversely, for each $\mathbf{x} = (x_\omega^k)_{\omega,k}$ satisfying $x_\omega^k \in X_{\omega,k}$ for every ω and k, there exists an SP equilibrium σ with this threat configuration. Moreover, the payoff of each player $i \in N$ in that equilibrium σ equals ϕ^i of formula (12), where the probability distribution $Q \equiv Q^{\mathbf{x}}$ is determined by the collection $\mathbf{x} = (x_\omega^k)_{\omega,k}$ according to (10) and (11).*

Remarks. (1) The threats x_ω^k and the payoffs do not depend on ρ. Moreover, the determination of any set X_ω^k can be done independently of any other such set. This holds here, in the TU case, but not in general, where optimal threats are determined recursively (i.e., one needs to determine first the optimal threats and proposals at all states that correspond to the subgames of ω).

(2) In every state, the payoffs and proposals are determined in the same way, by considering only the appropriate subgame.

Proof of Proposition 7. Let σ be an SP equilibrium. For each $\omega = (S, b^{N\setminus S})$ and $k \in S$, let $g_\omega := u^S(\zeta^S_\omega, b^{N\setminus S})$ and $g_{\omega,k} := u^S(\zeta^S_{\omega,k}, b^{N\setminus S})$; since $g_{\omega,k}$ is individually rational and Pareto efficient in $C(\omega)$ (recall Lemma 2 (ii)), (6) implies that

$$\sum_{i\in S} g^i_{\omega,k} = v(\omega). \tag{13}$$

Therefore, by (1), the same holds for g_ω:

$$\sum_{i\in S} g^i_{\omega} = v(\omega). \tag{14}$$

Moreover, (6) implies that maximizing the k-th coordinate $g^k_{\omega,k}$ is equivalent to minimizing all the other coordinates $g^i_{\omega,k}$, and so conditions (3) are satisfied as equalities, i.e., for every $i \in S\backslash k$,

$$g^i_{\omega,k} = \rho g^i_\omega + (1-\rho) u^i(\eta^S_{\omega,k}(x^k_\omega)). \tag{15}$$

Summing this over $i \in S\backslash k$ yields

$$v(\omega) - g^k_{\omega,k} = \rho(v(\omega) - g^k_\omega) + (1-\rho)t \tag{16}$$

where

$$\begin{aligned} t &:= \sum_{b^k\in A^k} x^k_\omega(b^k) \sum_{i\in S\backslash k} u^i(\zeta^{S\backslash k}_{(\omega||b^k)}, (b^{N\setminus S}, b^k)) \\ &= \sum_{a^k\in A^k} x^k_\omega(b^k)\, v(S\backslash k, (b^{N\setminus S}, b^k)) \end{aligned} \tag{17}$$

(we have used (14) for $S\backslash k$). Rewrite (16) as

$$\begin{aligned}(1-\rho)(v(\omega)-t) &= g^k_{\omega,k}-\rho g^k_\omega \\ &= \left(1-\frac{\rho}{|S|}\right)g^k_{\omega,k}-\frac{\rho}{|S|}\sum_{j\in S\backslash k}g^k_{\omega,j}.\end{aligned} \tag{18}$$

Therefore, in order to maximize $g^k_{\omega,k}$ (i.e., to satisfy (4)), one must minimize t (the other terms are fixed here). But t depends only on the threat x^k_ω (and the given function v), and so $t=\tau_{\omega,k}$ and $x^k_\omega\in X^k_\omega$; therefore (see (9) and (18)):

$$g^k_{\omega,k}=\rho g^k_\omega+(1-\rho)D^k v(\omega). \tag{19}$$

Let

$$h^k_{\omega,i}:=u^k(\eta^S_{\omega,i}(x^i_\omega))=\sum_{b^i\in A^i}x^i_\omega(b^i)\,g^k_{(\omega||b^i)} \tag{20}$$

be the payoff of k when $i\neq k$ becomes inactive and his threat x^i_ω is implemented; then (by (15), interchanging i and k):

$$g^k_{\omega,i}=\rho g^k_\omega+(1-\rho)h^k_{\omega,i}.$$

Adding this over all $i\neq k$ together with equation (19) yields $|S|g^k_\omega=\rho|S|g^k_\omega+(1-\rho)\left(D^k v(\omega)+\sum_{i\in S\backslash k}h^k_{\omega,i}\right)$, or

$$g^k_\omega=\frac{1}{|S|}\left(D^k v(\omega)+\sum_{i\in S\backslash k}h^k_{\omega,i}\right). \tag{21}$$

Substituting (20) yields recursively formula (12). $\square$

At this point it is useful to analyze a simple example.

Example. Let $N=\{1,2,3\}$, and for each $i\in N$ put $A^i=\{0,1\}\times N$, with generic element $a^i=(c^i,d^i)$ where $c^i\in\{0,1\}$ and $d^i\in N$. The payoffs are as follows: $u^i(a)=u^i_1(c)+u^i_2(d)$, where: $u^3_1(c)=1$ when $c^1=c^2$ and $c^3=1$, and $u^3_1(c)=0$ otherwise; $u^i_1\equiv 0$ for $i=1,2$; and $u^i_2(d):=|\{j:d^j=i\}|-1$. Thus, according to the u^i_1-part of the payoff functions u^i, player 3 gets a payoff of

1 when he chooses $c^3 = 1$ and the c^i-s of players 1 and 2 match; all other payoffs are 0. The effect of the u^i_2-part is to allow transfers and therefore to make the game TU: $d^j = i$ means that player j transfers one unit to player i (i.e., the payoff of i increases by 1, and that of j decreases by 1; note that $d^j = j$ means that j makes no transfer).

Consider the PC procedure. The optimal threats are determined by (7); this implies that there will never be any transfers (since this only increases the worth of the remaining players after the proposer becomes inactive), and so we will only specify the c^i coordinate. Player 3 can always guarantee that the remaining players get 0 (by using the threat $c^3 = 0$). The only case where his marginal contribution is not $D^3 v(\omega) = 1 - 0 = 1$ is when the threats of the players that became inactive before him made v equal to 0; this happens only when 3 is last, and then the optimal threat of the second player to become inactive, say player 2, is the opposite of that of player 1 (i.e., $c^2 = 1 - c^1$). Therefore the SP equilibrium payoffs are $(1/6, 1/6, 2/3)$.

It is interesting to compare this to the more familiar approaches. The von Neumann–Morgenstern coalitional function is $v(N) = 1$, $v(1, 3) = 1/2$ (player 1 plays $c^1 = 0$ and $c^1 = 1$ with half-half probabilities—denote this $(1/2, 1/2)$—and player 3 plays $c^3 = 1$), $v(2, 3) = 1/2$, and $v(S) = 0$ otherwise. The Shapley value of this v is $(1/4, 1/4, 1/2)$.

When going to the Harsanyi coalitional function we get[6] $v(N) = 1$, $v(1) = 1/4$, $v(2, 3) = 3/4$, $v(2) = 1/4$, $v(1, 3) = 3/4$, $v(3) = 1/2$, $v(1, 2) = 1/2$. This is an inessential game, and its value is again $(1/4, 1/4, 1/2)$.

Interestingly, the SP equilibrium payoffs $(1/6, 1/6, 2/3)$ seem to reflect better the underlying situation. The payoffs $(1/4, 1/4, 1/2)$ are what one would expect if $\{1, 2\}$ acted as one player, and then split the payoff. But it seems natural that the need to coordinate between 1 and 2 has some cost to them, and the payoff vector $(1/6, 1/6, 2/3)$ captures this better.

[6] Take for example $\{1\}$ vs. $\{2, 3\}$. The optimal strategies are $(1/2, 1/2)$ for 1, vs. $(1/2, 1/2)$ for 2 and $c^3 = 1$ for 3, which give payoffs of 0 to $\{1\}$ and $1/2$ to $\{2, 3\}$. Therefore $v(1) = 0 + (1 - 0 - 1/2)/2 = 1/4$ and $v(23) = 1/2 + (1 - 0 - 1/2)/2 = 3/4$.

5 Equilibria with Fixed Threats

Threats are of the essence of the theory we are presenting in this paper. It is because of the strategic linkage across coalitions captured by them that, for example, we cannot in any general and meaningful sense factor our analysis through the coalitional forms of standard cooperative game theory: there is no "worth" of a coalition that is independent of the actions—the threats—of the players outside the coalition.

This difficulty at the foundations of cooperative game theory has, of course, been recognized for a long time. It has led, on the one hand, to the development of extensions of the notion of coalitional form (perhaps the most well known are the "games in partition form" of Thrall and Lucas, 1963; see Myerson, 1977, Maskin, 2003, de Clippel and Serrano, 2005, Macho-Stadler, Perez-Castrillo and Wettstein, 2007, for more recent work) and on the other to the consideration of particular situations where the classical form could be justified (for example the c-games of Shapley–Shubik, see Shubik, 1983, p.130).

Nonetheless, the discussion of the previous section, and especially expression (12) for the computation of the SP equilibria, suggests a close connection to the cooperative-game solutions related to the Shapley (1953) value. In this section we shall throw some light on this connection.

It is certainly the case that along an equilibrium path only the particular actions that may arise as threats matter. But even then the threat of a proposer may depend on the current set of active players and on the threats of the preceding proposers. Still, if the threats happen to be independent of the previous history, we could indeed associate a coalitional form to the particular equilibrium, and we could then analyze how the equilibrium payoffs relate to the cooperative game theory solutions of the coalitional form. This we shall now do.

Definition. Let G be a strategic game and σ an SP equilibrium of the PC procedure. For every player $k \in N$ let $f^k \in A^k$ be a pure action of k. We say that σ has *fixed threats* $(f^k)_{k \in N}$ if, with probability 1 (that is, along the equilibrium path), whenever k is the proposer then the announced threat is

f^k.

Observe that the definition does not put any restriction on threats off the equilibrium path.

Next, given a strategic game G and an SP equilibrium σ with fixed threats $(f^k)_{k\in N}$, we say that the NTU coalitional game $(N, V_{G,\sigma})$ is *derived* from G and σ if

$$V_{G,\sigma}(S) = \{c \in \mathbb{R}^S : c \leq u^S(z^S, f^{N\setminus S}) \text{ for some } z^S \in \Delta(A^S)\}$$

for every coalition $S \subset N$. We have:

Proposition 8 *Let $(N, V_{G,\sigma})$ be a game that is derived from the strategic game G and the fixed-threat equilibrium σ. Suppose that $(N, V_{G,\sigma})$ is a TU game in the individually rational region.*[7] *Then the payoffs induced by σ equal the Shapley values of $(N, V_{G,\sigma})$ and its subgames. Moreover, if ρ is close to* 1*, then the payoffs of the proposals made by the different players will also be close to the Shapley values of $(N, V_{G,\sigma})$.*

Thus, when $(N, V_{G,\sigma})$ is a TU coalitional game—let $v \equiv v_{G,\sigma}$ denote its TU coalitional function—the outcome configuration $\boldsymbol{\zeta}$ of σ satisfies $u^i(\zeta^S_{(S,f^{N\setminus S})}) = \mathrm{Sh}^i(S, v_{G,\sigma})$ for every $i \in S \subset N$; moreover, as $\rho \to 1$ we also have $u^i(\zeta^S_{(S,f^{N\setminus S}),k}) = \mathrm{Sh}^i(S, v_{G,\sigma})$ for every $i, k \in S \subset N$.

Proof. Similar to the proof of Proposition 7; see in particular the explicit computational formula there. Note that the fixed threats imply that what a coalition can obtain is well defined, in the sense of not depending on the order in which the inactive players have dropped out. □

Proposition 8 does show that in a very natural sense the solution concept we develop in this paper, SP equilibrium of the PC procedure, is an extension to a larger context of the Shapley value solution for TU coalitional form games.

[7] I.e., for every $S \subset N$ there is a real number $v(S)$ such that $c \in V_{G,\sigma}(S)$ and $c^i \geq r^i$ for all $i \in S$ if and only if $\sum_{i\in S} c^i \leq v(S)$.

What happens in the general NTU case? One may conjecture, that, as in Hart and Mas-Colell (1996a), as ρ approaches 1 the SP equilibrium payoffs approach a Maschler–Owen (1992) consistent NTU value of $(N, V_{G,\sigma})$. It is not difficult to see that this is indeed the case if, for every S, the limit of the SP equilibrium payoffs of S lies in a smooth piece of the efficient boundary of[8] $V_{G,\sigma}(S)$. Since this set is a convex polytope, the condition amounts to the requirement that each limit lies in the interior of some $(|S|-1)$-dimensional face of the polytope. In particular, this will be automatically satisfied if $(N, V_{G,\sigma})$ is a hyperplane game (Maschler and Owen, 1989; of course only the individually rational region matters, as in Proposition 8 for the TU case). But a general analysis of the non-smooth case is needed.

6 Games with Damaging Actions

Are there classes of games in strategic form that, from the standpoint of the PC procedure, lend themselves to being summarized by means of the coalitional form of cooperative game theory? Presumably, these would be concrete specifications of the c-games of Shapley–Shubik.

In this section we exhibit one such class of games by presenting a property of strategic forms that, for TU games, implies the existence of an SP equilibrium enjoying the fixed-threat property. The NTU case is discussed in the next section.

It is reasonable to expect that the strategic linkage through threats is bound to be simpler in situations where there is some form of "strategic dominance" or "universality" in the threats used by players. This suggests the following:

Definition. Given a game G, a player $k \in N$ has a *damaging action* $d^k \in A^k$ if $u^i(d^k, a^{N \setminus k}) \leq u^i(a)$ for every action profile $a \in A$ and every player $i \neq k$. A game G is a *d-game* if every player $k \in N$ has a damaging action.

[8] See Hart and Mas-Colell 1996a, Proof of Proposition 8 (with the correction at http://www.ma.huji.ac.il/hart/abs/nbarg.html) for the reason for the smoothness requirement.

That is, a d-game is such that whatever the play is, if player k switches his action to d^k then the payoffs of all the other players decrease or stay the same; it is a strong property. The next proposition shows that, indeed, the "d" concept helps relate our approach to cooperative game theory.

Proposition 9 *Let G be a strategic TU game. Suppose that G is a d-game. Then there exists a fixed-threat SP equilibrium of the PC procedure where each player i uses a damaging action as threat.*

Proof. Let d^k be a damaging action of player k. Recall (Proposition 7) that at an SP equilibrium of a strategic TU game a proposer k chooses a threat $x^k \in \Delta(A^k)$ at state $\omega = (S, b^{N\backslash S})$ so as to minimize $v(S\backslash k, (b^{N\backslash S}, x^k))$, the sum of the payoffs of the remaining players if the proposer becomes inactive. Obviously, the pure threat d^k will do the job for k, at any state. □

7 Market Games Are Not c-Games

Propositions 8 and 9 highlight in a clear way the relationship between the bargaining theory we develop in this paper and classical cooperative game theory: if threats are "self-evident" then they can be taken as fixed threats, a coalitional form emerges in the obvious manner, and the analysis can proceed by taking the coalitional form as the basic datum and appealing to the extensive and rich theory of cooperative games. But Proposition 9 was stated for the TU case. In this section we shall see by means of an example that the result is no longer true for the general NTU situation and that this is so for entirely non-pathological reasons, that is, for reasons that seem inherent in the nature of strategic bargaining among the many. It is therefore very questionable whether, even under the strong hypothesis of the players having damaging actions, bargaining theory in the strategic form can justifiably be factored through cooperative game theory (except in the TU case).

The example will be built over a pure exchange economy satisfying the standard conditions (no externalities, concavity, and monotonicity of preferences, etc.). We choose this framework because exchange economies have been thought to be the paradigmatic cases of c-games ("c" stands for "consent"), i.e., the sort of games where the self-evident coalitional form was fully adequate (in the interesting discussion of Shubik, 1983, p. 131, it is said "in economic theory, games satisfying the consent condition arise in many places, most notably in models of pure competition without externalities"). We shall see in the example below that the obvious damaging threat of never sharing your endowment is not always the optimal threat! The phenomenon is related to the well-known transfer and endowments paradoxes of general equilibrium theory (see, for example, Postlewaite, 1979; also, Mas-Colell, 1976), but we should emphasize that here these emerge internally to the theory, i.e., within well-specified rules of a game.

Example. An exchange economy with 4 commodities and 3 traders. Let the commodities be b, c, f, g, and the traders, $1, 2, 3$. The initial endowments are

$$\begin{aligned} e^1 &= (0,0,1,1), \\ e^2 &= (0,1,0,0), \\ e^3 &= (1,0,0,0), \end{aligned}$$

and the utility functions are

$$\begin{aligned} u^1(b,c,f,g) &= b, \\ u^2(b,c,f,g) &= b+c-1, \\ u^3(b,c,f,g) &= \frac{1}{2}c + \max_{\substack{b'+b''=b \\ b',b''\geq 0}} \left\{ \frac{1}{2}\min\{b',f\} + \min\{b'',g\} \right\}. \end{aligned}$$

The goods b and c are mediums of exchange ("money"); player 2 has a "technology" which takes b as input and transforms it into "utils" subject to capacity constraints determined by f and g, where the productivity through g is twice as high as the one though f.

We make the exchange economy into a strategic game in a natural way, as first formally suggested by Scarf (1971): each player i distributes his endowment e^i among the 3 players: $e^i = \sum_{j=1}^{3} d^{i,j}$, where $d^{i,j} \in \mathbb{R}^4_+$ is the bundle transferred from i to j; the outcome (final holding) of player j is thus $h^j = \sum_{i=1}^{3} d^{i,j}$, and his payoff is $w^j = u^j(h^j)$.

Note that in this game every player has a damaging action: to keep all his endowment for himself (i.e., $d^{i,i} = e^i$). Suppose first that these are indeed the threats; the resulting coalitional function is (in the individually rational region, i.e., where all payoffs are nonnegative) a TU game for all coalitions except $\{2,3\}$. We get $v(i) = 0$ for all i, $v(1,2) = 0$, and $v(1,3) = v(1,2,3) = 1$. As for $\{2,3\}$, the Pareto efficient boundary is on the line $w^2 + 2w^3 = 1$. Computing the payoff vectors yields[9] $(0,0,\cdot)$ for $\{1,2\}$, $(1/2,\cdot,1/2)$ for $\{1,3\}$, and $(\cdot,1/2,1/4)$ for $\{2,3\}$; extending them to efficient payoff vectors for the grand coalition $N = \{1,2,3\}$ and then averaging gives the final outcome of[10] $(1/4, 1/6, 7/12)$.

However, this does not yield an equilibrium, because player 1 has a better threat when he is the proposer in the grand coalition, namely, to transfer his unit of the f good to player 3. Notice that this threat does not change the nonnegative attainable set for coalition $\{2,3\}$, but player 3 now gets by himself 1/2 rather than 0. The negotiating terms in coalition $\{2,3\}$ have been altered, and the outcome of this coalition becomes $(\cdot, 0, 1/2)$; this implies that player 1 can make a demand of 1/2 (instead of the 1/4 that he could ask for when the threat was to keep his own resources). The outcome of the grand coalition is now $(1/3, 0, 2/3)$—and so player 1's payoff has increased from 1/4 to 1/3 by the above deviation. In a sense, by this action player 1 has successfully manipulated in his favor the bargaining between 2 and 3 (note that this could not happen in the TU case, because only the *total* payoff of $\{2,3\}$ matters to 1).

So, what are the SP equilibria in this example? Since the efficient boundaries of the attainable sets for $\{1,2,3\}$ and all coalitions except $\{2,3\}$ are TU, keeping one's endowment is optimal, except for player 1 in the grand coalition. His threat in this case must minimize the sum of the payoffs of 2 and 3 in the subgame after 1 becomes inactive (since 1 gets the difference between $v(1,2,3) = 1$ and that sum). Now this sum is at least 1/2, since

[9] A dot $(\cdot)$ is put for the coordinate of the missing player.

[10] See Hart (2004, Section 5) for a similar computation.

the outcome $(\cdot, 0, 1/2)$ is always feasible for $\{2, 3\}$ (even without any transfers). By transferring 1 unit of good f to player 3, player 1 makes sure that $(\cdot, 0, 1/2)$ is necessarily the outcome of $\{2, 3\}$—so this is the optimal threat of player 1 in coalition $\{1, 2, 3\}$. Thus the unique SP equilibrium payoffs for the grand coalition are $(1/3, 0, 2/3)$ as seen above. Note in particular that the unique optimal threat of 1 in $\{1, 3\}$ is to keep his own endowment, and so the SP equilibria[11] do not have the fixed-threat property.[12]

In conclusion, strategic market games, a classical instance of the so-called c-games, are not really c-games: one cannot simply define the coalitional function as what a coalition can do with the total endowment of its members (except in the TU case). Our point, however, is more general. In the example above player 1, by using a suitable threat (which is not a damaging threat), can alter—to his advantage—the relative bargaining powers of players 2 and 3 in the subsequent negotiation. In the general NTU case, where the specific subsequent agreement of a coalition of players matters to the proposer (whereas in the TU case only the sum of payoffs matters to him), this is bound to be pervasive.

From a different perspective, we can view the analysis just made as underlying the existence of a substantial theoretical gap between the TU and the NTU situations. One cannot take for granted that the interesting phenomena that may hold for the former will carry over to the latter (for a different question—the equivalence principle—we made a similar point in Hart and Mas-Colell, 1996b).

8 Extensions

We mention here a number of possible extensions, generalizations, and questions suggested by this study:

[11] While the SP equilibrium outcomes are unique, the strategies are not (for example, player 1's threat in coalition in coalition $\{1, 2\}$ is arbitrary).

[12] The curious reader might wonder what are the Walrasian equilibrium payoffs in this example. They are $(0, 0, 1)$, which arise from zero prices for goods f and g, and a positive price for b that is no more than twice the price for c.

(a) The PC procedure may be modified in various ways. Two options that appear interesting are:

(i) A threat $x^k \in \Delta(A^k)$ is not realized immediately after a rejected proposer k becomes inactive, but rather at the end of the procedure. Thus a state consists of the set of active players S together with the fixed mixed actions of the inactive players $(x^i)_{i \in N \setminus S} \in \prod_{i \in N \setminus S} \Delta(A^i)$.

(ii) Dispose with the threats altogether and make the inactive players lose their power to choose their actions; thus a proposal is now a $z^N \in \Delta(A^N)$, but only the active players are asked to accept it.

(b) Propose and study bargaining procedures that correspond to the Harsanyi N-person generalization of Nash's two-person variable-threat game.

(c) Characterize situations where fixed threats and damaging actions obtain.

(d) Characterize exchange economies where keeping one's endowment is an optimal threat, and study the connections to other solution concepts.

References

Aumann, R. J. (1959), Acceptable Points in General Cooperative n-Person Games, in *Contributions to the Theory of Games IV, Annals of Mathematics Study 40*, A. W. Tucker and R. D. Luce (eds.), Princeton: Princeton University Press, 287–324.

Binmore, K., M. Osborne, and A. Rubinstein (1992), "Noncooperative Models of Bargaining," in *Handbook of Game Theory, with Economic Applications*, Volume I, R. J. Aumann and S. Hart (eds.), Amsterdam: Elsevier North-Holland, 179–225.

de Clippel, G. and R. Serrano (2005), "Marginal Contributions and Externalities in the Value," Brown University (mimeo).

Harsanyi, J. C. (1959), "A Bargaining Model for the Cooperative n-Person Game, in *Contributions to the Theory of Games IV, Annals of Mathematics Study 40*, A. W. Tucker and D. Luce (eds.), Princeton: Princeton University Press.

Harsanyi, J. C. (1963), "A Simplified Bargaining Model for the n-Person Cooperative Game," *International Economic Review* 4, 194–220.

Hart, S. (2004), "A Comparison of Non-Transferable Utility Values," *Theory and Decision* 56, 35–46.

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

Hart, S. and A. Mas-Colell (1996b), "Harsanyi Values of Large Economies: Non-Equivalence to Competitive Equilibria," *Games and Economic Behavior* 13, 74–99.

Hildenbrand, W. (1974), *Core and Equilibria of a Large Economy*, Princeton: Princeton University Press.

Houba, H. and W. Bolt (2002), *Credible Threats in Negotiations: A Game-Theoretic Approach*, Boston: Kluwer.

Macho-Stadler, I., D. Pérez-Castrillo and D. Wettstein (2007), "Sharing the Surplus: An Extension of the Shapley Value for Environments with Externalities," *Journal of Economic Theory* 135, 339–356.

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 *Rational Interaction*, R. Selten (ed.), Berlin: Springer-Verlag, 5–12.

Mas-Colell, A. (1976), "En Torno a una Propriedad Poco Atractiva del Equilibrio Competitivo," *Moneda y Credito* 136, 11–27.

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

Myerson, R. B. (1977), "Value of Games in Partition Function Form," *International Journal of Game Theory* 6, 23–31.

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

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

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

Osborne, M. and A. Rubinstein (1994), *A Course on Game Theory*, Cambridge, MA: The MIT Press.

Postlewaite, A. (1979), "Manipulation via Endowments," *Review of Economic Studies* 46, 255–262.

Scarf, H. (1971), "On the Existence of a Cooperative Solution for a General Class of N-Person Games, *Journal of Economic Theory* 3, 169–181.

Shapley, L. S. (1953), "A Value for n-Person Games," in *Contributions to the Theory of Games II, Annals of Mathematics Studies* 28, H. W. Kuhn and A. W. Tucker (eds.), Princeton: Princeton University Press, 307–317.

Shapley, L. S. (1969), "Utility Comparison and the Theory of Games," in *La Décision*, Paris: Editions du CNRS, 251–263.

Shubik, M. (1983), *Game Theory in the Social Sciences*, Cambridge, MA: MIT Press.

Thrall, R. M. and W. F. Lucas (1963), "n-Person Games in Partition Function Form," *Naval Research Logistics Quarterly* 10, 281–298.

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