

# Unconditional threats and the core<sup>\*</sup>

József Sákovics<sup>†</sup>

Universitat de les Illes Balears

September 2026

## Abstract

The core's standard definition allows a blocking coalition to divide its joint payoff in whatever way the challenged allocation makes convenient, so that a coalition may threaten with different agreements against different proposals. We require instead that a coalition's threat be *unconditional*: the outcome it expects from the subgame it would play on its own. The consequences are economic as much as technical. We exhibit agreements in which a player who contributes nothing to any coalition is nevertheless paid, and in which players who would otherwise compete for a partner gain by declining to compete, neither being any longer undercut by an agreement tailored to the deal on the table. The resulting solution, the ut-core, contains the core, and in cases of the above kind is non-empty where the core is silent. The construction turns any solution concept for the subgames into another one for the game itself, the ut-core being the image of the core; the solution that reproduces itself under the operation is its recursive counterpart, not nested with it in either direction. We characterize the solution for three-player games, and also establish non-emptiness for symmetric superadditive games.

---

<sup>\*</sup>I thank Roberto Burguet, Françoise Forges, Estela Sánchez-Rodríguez, Tamás Solymosi, and Ina Taneva for useful comments. I am grateful for financial assistance from grant PID2020-115018RB-C33 funded by MICIU/AEI /10.13039/501100011033.

<sup>†</sup>j.sakovics@uib.eu

**Keywords:** core extension; blocking; credibility; unconditional threats; core center; transferable utility

**JEL classification:** C71, C78

## 1 Introduction

A central feature of cooperative games is that players form coalitions and enter into binding agreements on how to distribute the surplus generated. With that description in mind we reconsider the core (Gillies, 1953, 1959), which was not originally defined axiomatically – although axiomatizations have since been provided (see, for example, Peleg, 1992) – but reflects a stability requirement first pointed out by Edgeworth (1881): no subcoalition should be able to secure an agreement that makes all its members strictly better off. The standard reading lets a coalition block whenever its worth exceeds the total putatively assigned to its members, which leaves it free to divide that worth differently against each allocation it attacks. We require instead that the agreement a subcoalition would reach be *unconditional*: the same whatever is being blocked.

The idea is best seen in an example. Suppose Ann, Bob and Cai face a “two-person job”: individually they generate nothing, any pair generates 8, and so does the grand coalition. The core is empty, since no division of 8 among the three can leave every pair with 8. But that conclusion relies on letting a blocking coalition tailor its agreement to the allocation it attacks. Against  $(0, 4, 4)$ , Ann and Bob can block only by agreeing on  $4 - \varepsilon$  for Ann; against  $(4, 0, 4)$  the same pair must agree on  $4 + \varepsilon$  for her. One coalition, two different “binding agreements”, chosen according to what is on the table. This fits badly the idea that these agreements are determined by the coalition’s own opportunities.

A complementary argument, closer in spirit to non-cooperative reasoning but still short of a non-cooperative analysis, reinforces the concern. In a negotiation among many players, with many possible blocking coalitions, the idea that a subcoalition can

form and commit to a side deal *during* the negotiation in the grand coalition seems suspect. It is more natural to suppose that a subcoalition must first leave before *committing* to internal agreements. It cannot then threaten with an arbitrary side agreement, which could be renegotiated once the split occurs. Its threat is simply to exit, and it blocks *credibly* only if the expected outcome of the ensuing subgame makes its members better off.<sup>1</sup>

The two requirements are distinct, and they do different work. *Credibility* asks that a threat be something the coalition could actually obtain on its own. Being *unconditional* asks that it not depend on the agreement being attacked. Credibility alone turns out to leave the core unchanged; what binds is the combination, and both are delivered at once by the same device, namely taking the threat to be what the coalition expects from parting ways with the grand coalition.

Section 4 develops the solution in stages. Restricting blocking threats to the core of the corresponding subgame leaves the core unchanged, strengthening the invariance result of Ray (1989); what binds is requiring each coalition to rely on a fixed element of its subgame core whenever it blocks. In our leading case that element is the center of gravity of the core – the expected allocation under a uniform prior, first proposed and analyzed by González-Díaz and Sánchez-Rodríguez (2007) – and we call the result the *unconditional-threat core*, or ut-core.

We then isolate what the construction requires: a solution concept, which delivers the set of outcomes a departing coalition can look forward to, and a belief over that set, which determines which of them it expects. Threat and solution are thereby distinct objects: a departing coalition forms an expectation over an outcome the analysis leaves open, whereas the grand coalition, having nowhere to depart to, is predicted by the set itself. The construction turns any solution concept into another one, with the ut-core the image of the core. We can take the requirement one step further. The core is not a

---

<sup>1</sup>While this second interpretation appeals to sequential rationality, it does so outside any underlying non-cooperative game. The comparison with subgame perfection is instructive. Subgame perfection eliminates threats of actions that are not sequentially rational but whose presence sustains some equilibrium outcomes, and thereby *refines* the set of Nash equilibria. We eliminate non-credible threats that block non-core outcomes, and thereby *enlarge* the core.

fixed point of the operation, so the ut-core retains classical threats while abandoning a classical solution; asking instead for a solution concept that generates itself yields the recursive solution, which we show to be the unique such fixed point, to coincide with the ut-core for three players, and to be nested with it in neither direction.

Every solution in the family contains the core, so it is non-empty wherever the core is – and, as we show, in classes of games where the core is not. The question is therefore not whether these solutions do better than the core on this score, since they cannot do worse, but how much further they reach. We answer this for symmetric superadditive games, for superadditive games with at most three players – where we also characterize the solution completely, as a union of four convex polygons – and for a class in which only pairs are able to block. A sequence of examples then highlights the features of the solution that matter for economic analysis.

## 2 Relation to the literature

Ray (1989) is our closest precursor. He too asks whether blocking is credible, but qualifies the coalition rather than the side agreement: a coalition is credible if its members can reach some agreement immune to blocking by credible subcoalitions, and he shows that restricting attention to blocking by such coalitions leaves the core unchanged. In our view, the set-valued nature of the core reflects indeterminacy rather than discretion: it leaves open which element materializes, and does not thereby license the coalition to commit to one among them at will. Consequently, it is the (blocking) agreement that needs to be credible.

Burguet and Caminal (2020) ask the consistency question of a different solution concept. Starting from the Nash Bargaining Solution they ask how disagreement points are determined, arguing that it is incoherent for different players to base their threats on distinct counterfactuals. Similarly to our construction, their Solution with Consistent Outside Options requires all players to share the expectation about disagreement within any coalition. It is probabilistic, with several coalitions forming with positive

probability,<sup>2</sup> and, surprisingly, generically unique in superadditive games.

The aspiration literature (Cross, 1967; Selten, 1981; Bennett, 1983) is motivated by a similar concern: each player names a single price for their participation, and the coalitions that can afford those prices are the ones predicted to form. Aspiration sets are non-empty and have been used to extend the core (Bennett and Zame, 1988). The difference is where the requirement bites. There the *player* is held to one number across coalitions, and that number is part of what the solution determines; here the *coalition* is held to one division across the agreements it might block, fixed by the subgame it would depart into.

Our recursive solution shares its motivation with two literatures, but not its setting. In Chwe's (1994) largest consistent set a coalition deviates only if every chain of deviations it might set off leaves its members better off, so stability is a fixed point of the deviation process. Ray and Vohra's (1997) equilibrium binding agreements are recursive in the same spirit, a blocking coalition anticipating what the remaining players would then write among themselves. Both are more ambitious than what we do: they start from a strategic-form game and track the complementary players, whereas we stay within the characteristic function and ask only what a departing coalition expects of its own subgame.

A different tradition responds not to what a threat should be, but to the fact that the core is often empty. Aumann and Maschler (1964) let a player excluded from a blocking coalition answer an objection with a counter-objection, which neutralizes it. The resulting bargaining set is non-empty, and the kernel and the nucleolus (Schmeidler, 1969) descend from the same idea.<sup>3</sup> What influences outcomes there is the players left out of a coalition, which is closer in spirit to coalition-formation bargaining. Dutta et al. (1989) apply the same test to every objection in a chain; their "consistent" bargaining set contains the core, is contained in the bargaining set, and may be empty.

More directly still, one may weaken the blocking condition itself. The  $\varepsilon$ -core of

---

<sup>2</sup>It is sometimes deterministic, but necessarily probabilistic when the core is empty.

<sup>3</sup>Strictly speaking theirs is the pre-bargaining set, individual rationality having to be imposed separately.

Shapley and Shubik (1966) subtracts a uniform penalty from every coalition's worth, and the least core (Maschler, Peleg and Shapley, 1979) takes the smallest penalty for which the resulting set is non-empty. It is always non-empty and convex, and is computed by a linear program, none of which we can claim. While we extend the core as well, our goal is not non-emptiness, rather to investigate the implications of unconditional threats.

A further strand implements the core non-cooperatively, in the spirit of the Nash program, from Kalai et al. (1979) onwards; Platt (2009) surveys it. It has pursued concerns close to ours, such as sequential rationality and stationarity, but we build no non-cooperative game, and rather than seeking a scaffold for the core, we reinterpret credible blocking and follow that wherever it leads. Okada and Winter (2002) deserve special mention, since rather than specifying a game they characterize by axioms a class of games that implement the core. One axiom is subgame consistency (Harsanyi and Selten, 1988), requiring identical behavior across isomorphic subgames, which is close to our own requirement. That it nonetheless yields a different solution shows how far apart the cooperative and non-cooperative readings of such a restriction can be: ours excludes agreements, theirs sustains them.

### 3 Preliminaries

We denote a cooperative game with transferable utility by  $(N, v)$ , where  $N$  is the set of players indexed by  $1, 2, \dots, n$ , and the *characteristic function*,  $v(\cdot) : 2^N \setminus \emptyset \rightarrow \mathbb{R}$ , denotes the (aggregate) utility/payoff that each subset of players can obtain on their own, without specifying how it might be distributed among the members of the subset.<sup>4</sup>

A non-empty subset of the players,  $S \subseteq N$ , is referred to as a *coalition* ( $N$  is the *grand coalition*) and it has cardinality  $s \leq n$ . A game is *superadditive* if for any two disjoint coalitions,  $T$  and  $U$ ,  $v(T \cup U) \geq v(T) + v(U)$ . A game is *symmetric* if the value of a coalition only depends on the number of players forming it. For symmetric

---

<sup>4</sup>Transferability of utility presupposes that all the players have a quasi-linear utility function (in the good that is physically distributed) with the same slope (but possibly different intercepts).

games, we denote the value of any  $k$ -player coalition by  $v_{\#}(k)$ . A game is *convex* if the value added by any player is increasing in the size of the coalition it joins: for all  $i \in N$ , and  $T, U \subset \{N \setminus i\}$  with  $T \subset U$  we have that  $v(U \cup i) - v(U) \geq v(T \cup i) - v(T)$ .

For an arbitrary payoff vector,  $x \in \mathbb{R}^N$ , we define the *assignment* to coalition  $S$ ,  $x(S)$ , as the sum of the payoffs of the players in  $S$ :  $x(S) = \sum_{i \in S} x_i$  for all  $S \in 2^N \setminus \emptyset$ . A payoff vector is *individually rational* (IR), if and only if  $x_i \geq v(i)$  for every player  $i \in N$ . A payoff vector is *coalitionally rational* (CR), if and only if  $x(S) \geq v(S)$  for every coalition  $S \in 2^N \setminus \emptyset$ . Note that CR implies IR.

A *solution concept*  $\Phi$  assigns to every game a (possibly empty) set of individually rational payoff vectors distributing the worth of that game's grand coalition, a set that is therefore bounded; we call it the solution of the game. The *core* is the solution concept comprising all the payoff vectors that distribute the grand coalition's aggregate worth,  $v(N)$ , in a CR way (if there are any: otherwise the core is "empty"):

$$\text{core}(N, v) = \{x \in \mathbb{R}^N : x(N) = v(N) \text{ and } x(S) \geq v(S) \text{ for all } S \in 2^N \setminus \emptyset\}.$$

The *core center* of  $(N, v)$ , denoted by  $\mu(N, v)$ , assigns to each player their expected utility calculated over their payoffs in  $\text{core}(N, v)$ , using a uniform prior. In other words, it is the center of gravity of the core. If the core is empty, the core center does not exist.

Finally, for any coalition  $T \subset N$ , the *subgame* restricted to  $T$ ,  $(T, v_T)$ , is defined by  $v_T(S) = v(S)$  for all  $S \in 2^T \setminus \emptyset$ . Similarly, a payoff vector for the game restricted to  $T$  is denoted by  $x_T \in \mathbb{R}^T$ . A payoff vector,  $x_S \in \mathbb{R}^S$ , *strictly dominates* another,  $y_S \in \mathbb{R}^S$ , denoted by  $x_S > y_S$ , if and only if  $x_i > y_i$  for all  $i \in S \subseteq N$ . Accordingly, we write  $x_S \not\prec y_S$  for the negation of  $y_S > x_S$ , that is, whenever  $x_i \geq y_i$  for at least one  $i \in S$ .

For simplicity, in our examples we will denote the characteristic function of a game by a vector  $v \in \mathbb{R}^{2^n - 1}$ , depicting the payoffs of coalitions in increasing order of cardinality and indices. In particular, for a 3-person game

$$v = (v(1), v(2), v(3); v(\{1, 2\}), v(\{1, 3\}), v(\{2, 3\}); v(\{1, 2, 3\})) \in \mathbb{R}^7.$$

## 4 The solution with unconditional threats

This section proceeds in stages. Section 4.1 imposes credibility alone, requiring a blocking threat to lie in the core of the subgame played by the blocking coalition; this is stronger than the credibility requirement behind Debraj Ray's (1989) modified core, yet we show it still leaves the core unchanged. Section 4.2 adds unconditionality, which yields the ut-core, and then isolates the general construction behind it: a solution concept and a belief over it determine a threat map, and every threat map determines a solution. Section 4.3 applies that construction to itself. Every solution the construction generates contains the core, hence is non-empty whenever the core is; Section 4.4 asks how much further each of them reaches, and exhibits classes of games in which they are non-empty and the core is not.

### 4.1 Credible blocking

Our notion of credibility requires that any side agreement used to block a putative payoff vector belong to the solution of the subgame in which the blocking coalition is on its own.<sup>5</sup> As our starting point is the core, the most straightforward way to achieve credibility is to ensure that the threat belongs to the core of the subgame following the splitting away of the putative blocking coalition.<sup>6</sup> Let us state this formally.

**Definition 1** *A blocking threat  $x_T \in \mathbb{R}^T$  by coalition  $T$  is credible if and only if  $x_T \in \text{core}(T, v_T)$ .*

Ray (1989) qualifies coalitions, rather than coalitional agreements, as credible: singletons are credible, and a larger coalition is credible if some distribution of its worth is immune to blocking by credible subcoalitions. His condition therefore asks only

---

<sup>5</sup>In this context, credibility simply means that the payoff is attainable by the coalition on its own, not that it is necessarily preferable to available alternatives.

<sup>6</sup>Note that this approach implies that if the core of a subgame is empty, the corresponding coalition is unable to block at all. This property of the solution can be rationalized by the fact that an empty core signals the inability of the coalition to reach an agreement.

that the blocking coalition be able to agree on something, not that the agreement it blocks will be one of those. The three restrictions are nested. Immunity to blocking by credible subcoalitions is weaker than immunity to blocking by all of them, so every coalition with a non-empty core is credible in Ray's sense: letting only coalitions with a non-empty core block is therefore more restrictive than his condition, and Definition 1, which constrains the agreement and not merely the coalition, is more restrictive still.<sup>7</sup> Since admitting fewer blocks can only enlarge the set of unblocked assignments, the three solutions are ordered by inclusion, and all of them contain the core. It therefore suffices to show that the strongest of the three, ours, leaves the core unchanged; Ray's invariance result then follows.

**Proposition 1** *Restricting blocking threats to be credible leaves the set of unblocked assignments, the core, unchanged.*

**Proof.** Restricting the admissible threats can only weakly enlarge the solution set, so it suffices to show that every  $x \notin \text{core}(N, v)$  is still blocked. Among the subcoalitions with  $x(S) < v(S)$ , of which there is at least one, let  $T$  be one of minimal cardinality; every  $W \subset T$  then satisfies  $x(W) \geq v(W)$ . Split the deficit equally, setting

$$y_T = x_T + \frac{v(T) - x(T)}{|T|} \mathbf{1},$$

which strictly dominates  $x_T$  because  $x(T) < v(T)$ . Then  $y_T(T) = v(T)$  and  $y_T(W) \geq x(W) \geq v(W)$  for every  $W \subset T$ , so  $y_T \in \text{core}(T, v_T)$  and is therefore a credible threat.<sup>8</sup>

■

That the core survives this restriction may suggest that it is immune to credibility concerns. We wish to argue that this conclusion is unwarranted. We claim that in assessing the credibility of a coalition's blocking threat, the relevant object is the specific threat itself and the appropriate question is not whether the threatened outcome *could* arise in the subgame (that is, whether it lies in the core), but whether it *would*

---

<sup>7</sup>In principle, it could happen that a putative assignment is only blocked by a threat outside of the (non-empty) core.

<sup>8</sup>Imposing the restriction recursively changes nothing either: this  $y_T$  lies in the core of  $(T, v_T)$ , hence also in the larger set that is immune only to credible blocking.

arise – at least in expectation. Accordingly, we require that the (expected) outcome of negotiations within a potential blocking coalition should be used as a threat.

## 4.2 The unconditional-threat core

Making “the expected outcome” precise takes two ingredients: the set of agreements the coalition might reach, and a belief over that set.

**Definition 2** *A belief rule  $F$  assigns to every non-empty bounded set  $A$  a probability distribution  $F(A)$  supported on  $A$ .*

The belief rule is what a coalition’s expectation is formed by: it reads off a belief over whatever set of agreements it considers feasible. Our leading rule is the *uniform* one, the uniform distribution on  $A$ .<sup>9</sup>

**Definition 3** *Let  $\Phi$  be a solution concept and  $F$  a belief rule. The induced threat map  $\tau$  assigns to every  $S \in 2^N \setminus \{\emptyset, N\}$  with  $\Phi(S, v_S) \neq \emptyset$  the mean of  $F(\Phi(S, v_S))$ .*

We abbreviate  $\tau(S, v_S)$  to  $\tau(S)$  whenever the game is held fixed, and to  $\tau(S, v)$  where the game varies, the second argument entering only through  $v_S$ . We denote the domain of the threat map by

$$\mathcal{D}(\tau) = \{S \in 2^N \setminus \{\emptyset, N\} : \Phi(S, v_S) \neq \emptyset\}.$$

Any threat by  $S \in \mathcal{D}(\tau)$  is *efficient*:

$$\sum_{i \in S} \tau_i(S) = v(S), \tag{1}$$

because every element of  $\Phi(S, v_S)$  distributes  $v(S)$  and averaging preserves this. Moreover, it is *credible*, in the sense that it lies in the convex hull of  $\Phi(S, v_S)$ , because the

---

<sup>9</sup>Formally, normalized Hausdorff measure in the set’s own dimension: for Borel  $A$  with  $d = \dim_H A$  and  $0 < \mathcal{H}^d(A) < \infty$  we take  $F(A) = \mathcal{H}^d \upharpoonright_A / \mathcal{H}^d(A)$ , which covers every set arising below. Anonymity and covariance are immediate, Hausdorff measure being invariant under isometries and the normalization removing the scale factor.

mean of a distribution lies in the convex hull of its support. When  $\Phi(S, v_S)$  is convex, as the core is, credibility is the same as membership, and Definition 1 is recovered.

Efficiency also explains why large coalitions are weak in this framework. Since the threat of  $S$  distributes exactly  $v(S)$  across its members, some member must be offered no more than the per capita worth  $v(S)/|S|$ , and paying that one member defuses the block. A pair may thus cost  $v(S)/2$ , a triple at most  $v(S)/3$ , and so on: the larger the coalition, the cheaper it is to break. This is why, in all our examples, the binding constraints come from small coalitions.

Two invariance properties will matter. Let  $\pi$  be a permutation of  $N$ . Write  $\pi S = \{\pi(i) : i \in S\}$ , let  $\pi x$  be the payoff vector obtained by moving each player's coordinate with that player,  $(\pi x)_{\pi(i)} = x_i$ , and let  $\pi v$  be the game with  $(\pi v)(\pi S) = v(S)$  for every  $S$ . A threat map is *anonymous* if

$$\tau(\pi S, \pi v) = \pi \tau(S, v) \text{ for every } \pi, \text{ every } v \text{ and every } S \in \mathcal{D}(\tau),$$

so that relabeling the players relabels the threat, and relabels the domain with it.<sup>10</sup> For  $a \in \mathbb{R}^N$ , let  $v^a$  denote the *translate* of  $v$  by  $a$ , the game with  $v^a(S) = v(S) + \sum_{i \in S} a_i$ , obtained by raising each player  $i$ 's stand-alone worth by  $a_i$ . A threat map is *covariant* if its domain is the same for  $v$  and for  $v^a$  and  $\tau(S, v^a) = \tau(S, v) + a_S$  for every  $a$  and every  $S \in \mathcal{D}(\tau)$ , where  $a_S = (a_i)_{i \in S}$ : what a player is given outright is added to what they expect.<sup>11</sup>

A belief rule is asked for the same two:  $F$  is *anonymous* if  $F(\pi A)$  is the image of  $F(A)$  under  $\pi$ , and *covariant* if  $F(A + a_S)$  is its image under that translation. The uniform rule is both.

Our leading case takes  $\Phi$  to be the core and the belief to be uniform. The resulting threat is the core's center of gravity, the *core center*  $\mu$  (González-Díaz and Sánchez-

---

<sup>10</sup>The requirement is equivariance rather than invariance, and it is the coalitional form of the standard anonymity axiom: it asks that isomorphic subgames receive corresponding threats. It is stronger than symmetry, or equal treatment of equals, which constrains only how substitutable players within one game are treated.

<sup>11</sup>This is the translation half of covariance under strategic equivalence, the usual axiom also requiring  $\tau(S, \alpha v) = \alpha \tau(S, v)$  for  $\alpha > 0$ . The core center satisfies that as well, but no result below uses it.

Rodríguez, 2007; an axiomatic analysis of this rule is given in González-Díaz and Sánchez-Rodríguez, 2009), which is clearly anonymous and covariant.

**Definition 4** *A threat map is admissible if it is anonymous, covariant, and  $\mathcal{D}(\tau) \supseteq \mathcal{D}(\mu)$ .*

**Definition 5** *A blocking threat  $x_T \in \mathbb{R}^T$  by coalition  $T$  is credible and unconditional if and only if  $x_T = \tau(T)$ .*

**Definition 6** *A payoff vector  $x \in \mathbb{R}^N$  belongs to  $\text{ut-core}^\tau(N, v)$  if and only if it distributes the value of the grand coalition and, for every subcoalition in the domain of  $\tau$ , it is not strictly dominated for the members of that subcoalition by that subcoalition's threat:*

$$\text{ut-core}^\tau(N, v) = \{x \in \mathbb{R}^N : x(N) = v(N) \text{ and } x_S \not\prec \tau(S) \text{ for all } S \in \mathcal{D}(\tau)\}.$$

Equivalently,  $x$  survives blocking by  $S$  when some member of  $S$  is at least as well off under  $x$  as under the threat, so that

$$\text{ut-core}^\tau(N, v) = \{x : x(N) = v(N)\} \cap \bigcap_{S \in \mathcal{D}(\tau)} \bigcup_{i \in S} \{x : x_i \geq \tau_i(S)\}. \quad (2)$$

Singletons always lie in  $\mathcal{D}(\tau)$  for an admissible  $\tau$ , and (1) makes their threats  $v(i)$ , so every element is individually rational and the solution is bounded. Distributing the intersection over the unions in (2) then exhibits it as a finite union of polytopes, closed but in general not convex; Proposition 4 below carries the distribution out for three players.

The definition uses strict domination, and the choice matters in a way it does not for the core. A coalition blocking in the classical sense may choose its agreement freely, so asking for an agreement that is weakly better for all its members and strictly better for one is the same as asking for one that is strictly better for all: either is possible precisely when  $v(S) > x(S)$ . Our threat is not chosen but fixed at  $\tau(S)$ , and there the two readings differ. Which convention is right depends on the shape of the blocking coalition's opportunity set. When that set is the whole hyperplane of divisions of  $v(S)$

among the members of  $S$ , a member who would gain can slide surplus to one who is indifferent and remain feasible, so the two coincide; when it is curved, as in a game without side payments, he cannot, which is why the classical NTU core requires every member of a blocking coalition to gain strictly (Aumann, 1961). Fixing the agreement at  $\tau(S)$  is the extreme case of the same phenomenon, and we resolve it the same way: a gainer cannot buy off an indifferent member without abandoning the very agreement whose credibility is at issue. This is a restriction on the agreements available, not a denial of transferability.<sup>12</sup>

Every solution concept – via its threat map – generates another one.

**Definition 7** *Let  $\Phi$  be a solution concept,  $F$  a belief rule, and  $\tau^{\Phi, F}$  the threat map they induce. The generated solution  $R(\Phi, F)$  assigns to every game the ut-core built from that threat map:*

$$R(\Phi, F)(S, v_S) = \text{ut-core}^{\tau^{\Phi, F}}(S, v_S).$$

Writing ut-core for the leading case  $\tau = \mu$ , it is the image of the classical solution,

$$\text{ut-core} = R(\text{core}, \text{uniform}).$$

A simple example shows why our solution extends the core. Consider the three-player game  $v = (0, 0, 0; 8, 8, 8; 8)$ . Its core is empty, since ensuring that each two-player coalition earns at least 8 would take 12 units. At the same time  $(0, 4, 4)$ ,  $(4, 0, 4)$  and  $(4, 4, 0)$  are the only elements of the ut-core, since the subgame played by any pair is symmetric and so its threat is  $(4, 4)$ , so it suffices that no two players earn strictly less than 4 each. These allocations are not in the core, because any two players can agree on  $(4 - \varepsilon, 4 + \varepsilon)$  for some  $\varepsilon \in (0, 4)$ . That threat is credible in the sense of Definition 1 – it lies in  $\text{core}(\{i, j\}, v_{\{i, j\}})$  – but it is conditional on what is being blocked, since it is not the agreement the subgame is expected to produce.

The example shows that threats by non-singleton coalitions work differently from unilateral ones. The constraint that single players do not block – Individual Rationality

---

<sup>12</sup>The choice also shapes the solution: the sets removed above are open, so the ut-core is closed, which weak domination would destroy.

– places solutions in the quadrant above the autarky point. For unconditional threats the relation is flipped: we rule out only the solutions lying in the quadrant *below* the threat point, because to destroy credibility it is enough that one member of the coalition prefers not to block, so for blocking we need all of them to be willing. There is also the complication that there are many threat points, one for each coalition in the domain, rather than the single autarky point.

An implication of the two invariance properties will be used repeatedly.

**Lemma 1** *Let  $\tau$  be admissible and let  $S \in \mathcal{D}(\tau)$ . If the subgame  $(S, v_S)$  becomes symmetric once every player's stand-alone worth is normalized to zero, then  $\tau(S)$  awards  $(v(S) - \sum_{i \in S} v(i)) / |S|$  above autarky to each member of  $S$ .*

**Proof.** By covariance we may take  $v(i) = 0$  for every  $i \in S$ . The subgame is then invariant under every permutation of  $S$ , so by anonymity  $\tau(S)$  is invariant too, which forces all its coordinates to be equal; (1) fixes their common value at  $v(S)/|S|$ . ■

Efficiency and Lemma 1 between them determine  $\tau$  on singletons, on pairs, and on every subgame that is symmetric after normalization; all of these have a non-empty core, and so lie in every admissible domain.

We close this subsection by showing that every solution the construction generates is a genuine extension of the core.

**Proposition 2** *For every game  $(N, v)$  and every admissible threat map  $\tau$ ,  $\text{core}(N, v) \subseteq \text{ut-core}^\tau(N, v)$ .*

**Proof.** Let  $x \in \text{core}(N, v)$  and let  $S \subset N$  lie in  $\mathcal{D}(\tau)$ . By (1) and coalitional rationality,  $\sum_{i \in S} \tau_i(S) = v(S) \leq x(S)$ , so  $x_i \geq \tau_i(S)$  for at least one  $i \in S$ , that is,  $x_S \not\leq \tau(S)$ . As this holds for every such  $S$ , we have  $x \in \text{ut-core}^\tau(N, v)$ .<sup>13</sup> ■

---

<sup>13</sup>The proof uses only efficiency, so the inclusion in fact holds for every threat map, admissible or not.

### 4.3 The recursive solution

Despite delivering an unconditional threat, the construction of the previous section still suffers from an inconsistency: it posits a different solution concept for the original game than for its subgames. In this subsection we remove that weakness.

**Definition 8** *Let  $F$  be an anonymous, covariant belief rule. The  $r\text{-ut-core}^F$  is defined recursively by  $r\text{-ut-core}^F(\{i\}, v) = \{v(i)\}$  and, for  $|S| \geq 2$ ,*

$$\begin{aligned} & r\text{-ut-core}^F(S, v_S) \\ &= \left\{ x \in \mathbb{R}^S : x(S) = v(S) \text{ and } x_T \not\prec \rho^F(T) \text{ for all } T \subset S \text{ with } r\text{-ut-core}^F(T, v_T) \neq \emptyset \right\}, \end{aligned}$$

where  $\rho^F(T)$  is the mean of  $F(r\text{-ut-core}^F(T, v_T))$ .

Unlike the core, the set a coalition now forms its expectation over need not be convex, so the threat is credible only in the sense already noted: it is an actual expected value that lies in the convex hull of that set, but not necessarily in the set itself.

The definition implies that  $r\text{-ut-core}^F(N, v) = \text{ut-core}^{\rho^F}(N, v)$  for every game, that is, that the  $r\text{-ut-core}^F$  is a fixed point of the operation of Definition 7:

$$r\text{-ut-core}^F = R(r\text{-ut-core}^F, F), \quad (3)$$

so the recursion produces no new kind of object: it picks out the member of the  $\text{ut-core}^\tau$  family at the threat map it generates recursively. The uniform rule is our leading case, and we drop its superscript throughout, writing  $r\text{-ut-core}$  and  $\rho$ .

The next proposition establishes the  $r\text{-ut-core}^F$ 's main properties.

**Proposition 3** *Fix an anonymous, covariant belief rule  $F$ .*

(i) *The  $r\text{-ut-core}^F$  is well defined and  $\rho^F$  is admissible, with*

$$\mathcal{D}(\rho^F) = \{S : r\text{-ut-core}^F(S, v_S) \neq \emptyset\} \supseteq \mathcal{D}(\mu).$$

(ii) *The  $r\text{-ut-core}^F$  is the unique solution concept  $\Phi$  satisfying  $\Phi = R(\Phi, F)$ .*

(iii) *If  $|N| \leq 3$  then  $r\text{-ut-core}^F(N, v) = \text{ut-core}(N, v)$ .*

**Proof.** We induct on  $|S|$ , which serves both (i) and (ii). For  $|S| = 1$  the defining formula imposes no blocking constraint, so  $R(\Phi, F)(\{i\}, v) = \{v(i)\}$  whatever  $\Phi$  is, which starts the induction. Suppose  $\rho^F$  has been defined on every coalition with fewer than  $k$  members, and that any two fixed points of  $R(\cdot, F)$  agree on such coalitions. For  $|S| = k$  the defining formula determines  $\text{r-ut-core}^F(S, v_S)$ , and  $\rho^F(S)$  is defined exactly when that set is non-empty, which fixes  $\mathcal{D}(\rho^F)$ ; moreover  $R(\Phi, F)(S, v_S)$  depends on  $\Phi$  only through the threats at proper subcoalitions of  $S$ , so any two fixed points agree there as well. That the r-ut-core is itself a fixed point is its definition, read game by game.

That  $\rho^F(S)$  exists at each step follows from (2), applied to the threats already defined on the proper subcoalitions of  $S$ :  $\text{r-ut-core}^F(S, v_S)$  is a bounded finite union of polytopes, so  $F$  assigns a distribution to it and that distribution has a mean. For the uniform rule this is the center of gravity of the highest-dimensional part, which has finite, strictly positive measure in its own dimension.

For the rest of (i): every element of  $\text{r-ut-core}^F(S, v_S)$  distributes  $v(S)$ , hence so does any mean of a distribution over it, giving (1); the defining conditions commute with relabelings of the players and with translations of the characteristic function, so the set is anonymous and covariant, and so therefore is  $\rho^F$ ,  $F$  being anonymous and covariant itself; and since  $\text{r-ut-core}^F(S, v_S)$  is the solution generated by  $\rho^F$  on the proper subcoalitions of  $S$ , Proposition 2 gives  $\text{core}(S, v_S) \subseteq \text{r-ut-core}^F(S, v_S)$ , so every coalition with a non-empty core lies in  $\mathcal{D}(\rho^F)$ .

(iii) Every proper subcoalition then has at most two members, and admissibility fixes the threats of singletons and pairs whatever solution concept and belief rule generate them, by the remark following Lemma 1; so  $\rho^F$  and  $\mu$  agree throughout and the two solutions coincide. ■

Every result stated for an admissible threat map therefore holds for  $\rho^F$ , whichever anonymous, covariant rule  $F$  is used. From here on we work with the leading case.

The recursion differs from the ut-core through two channels, and they pull in opposite directions. The first is the domain. By part (i),  $\mathcal{D}(\rho) \supseteq \mathcal{D}(\mu)$ : a coalition

whose core is empty is silent under  $\mu$ , but threatens under  $\rho$  whenever its own r-ut-core is not empty. Were the two threats to agree wherever both are defined, this alone would give  $\text{r-ut-core}(N, v) \subseteq \text{ut-core}(N, v)$ . They do not agree, and that is the second channel. Both satisfy (1), so if  $\rho(S) \neq \mu(S)$  there are  $i, j \in S$  with  $\rho_i(S) > \mu_i(S)$  and  $\rho_j(S) < \mu_j(S)$ ; since the vectors that  $S$  blocks under a threat  $\tau$  form the box  $\{x \in \mathbb{R}^S : x_i < \tau_i(S) \text{ for all } i \in S\}$ , and two boxes whose defining vectors cross in this way are not nested, each threat blocks vectors the other does not. The smallest case is a three-player coalition sitting inside a larger game, where  $\rho(S)$  is the center of the ut-core of the subgame and  $\mu(S)$  that of its core; for such an  $S$  the two agree when its three two-player subcoalitions are all worthless, and we know of no other case in which they do. Coalitions of four or more may of course differ as well. It is not to be confused with part (iii) of the proposition, which compares the two solutions of a game that has only three players.

Neither channel dominates in general. Whether a crossing of threats reaches the grand coalition is a further question, since it depends on whether the vectors in question are attainable given feasibility and the restrictions imposed by the remaining coalitions. Example 4 in Section 6 settles it: neither solution contains the other.

## 4.4 How far we extend the core

We can generalize the three-player example  $v = (0, 0, 0; 8, 8, 8; 8)$  above to symmetric games with any number of players. For these, superadditivity alone turns out to be sufficient for non-emptiness.

Throughout we normalize the game so that every player is worth nothing on her own, writing  $v^0(S) = v(S) - \sum_{i \in S} v(i)$ . Covariance makes this free of charge:  $\text{ut-core}^\tau(N, v) = (v(i))_{i \in N} + \text{ut-core}^\tau(N, v^0)$  for every admissible  $\tau$ . Note that superadditivity implies  $v^0 \geq 0$ .

We need two lemmas before we can state our first existence result. For a symmetric game we write  $c_k = v_{\#}^0(k)/k$  for the *per capita worth* of a  $k$ -player coalition, and  $\hat{c}_k = \max_{m \leq k} c_m$  for its running maximum.

**Lemma 2** *Let  $(N, v)$  be symmetric and let  $|S| = k$ . Then  $\text{core}(S, v_S) \neq \emptyset$  if and only if  $c_k = \hat{c}_k$ , and in that case  $\mu(S, v_S)$  is the equal split, assigning  $v_{\#}(k)/k$  to every member of  $S$ ; equivalently  $c_k$  above autarky, the stand-alone worths being equal.*

**Proof.** Being invariant under relabelings and convex, the core is non-empty if and only if it contains the equal split: average the  $k!$  relabelings of any of its elements. The equal split is coalitionally rational if and only if  $mv_{\#}^0(k)/k \geq v_{\#}^0(m)$  for every  $m \leq k$ , that is, if and only if  $c_k \geq c_m$  for all  $m \leq k$ , which is precisely  $c_k = \hat{c}_k$ . ■

Observe that the maximum defining  $\hat{c}_k$  is attained at some  $m \leq k$  with  $c_m = \hat{c}_m$ , so by Lemma 2 the running maximum is always achieved by a coalition whose core is non-empty, and which therefore lies in the domain of every admissible threat map. This yields an exact criterion, independent of which admissible threat map is used.

**Lemma 3** *Let  $(N, v)$  be symmetric and let  $\tau$  be admissible. Then*

$$\text{ut-core}^{\tau}(N, v) \neq \emptyset \iff \hat{c}_{n-1} + \sum_{k=1}^{n-1} \hat{c}_k \leq v_{\#}^0(n). \quad (4)$$

**Proof.** By covariance we may take  $v(i) = 0$ , so that  $c_1 = \hat{c}_1 = 0$ . Every subgame of a symmetric game is symmetric, so by Lemma 1 the threat of a  $k$ -player coalition in the domain is the equal split,  $\tau(S) = (c_k, \dots, c_k)$ . Such a coalition blocks  $x$  exactly when each of its  $k$  members receives strictly less than  $c_k$ , which is possible if and only if at least  $k$  players do. Thus, ordering the payoffs decreasingly as  $x_{(1)} \geq \dots \geq x_{(n)}$ , the vector  $x$  survives blocking by  $k$ -player coalitions if and only if

$$x_{(n-k+1)} \geq c_k, \quad (5)$$

a constraint present for every  $k$  in the domain.

The cheapest decreasing vector satisfying all these constraints for  $k \leq n-1$  therefore sets  $x_{(j)}$  equal to the largest  $c_k$  over  $k \leq n-j+1$  in the domain, for  $j \geq 2$ , and  $x_{(1)} = x_{(2)}$ , as no constraint bears on the largest payoff beyond monotonicity. By the observation preceding the lemma that largest value is  $\hat{c}_{n-j+1}$ , whatever the admissible domain: enlarging the domain adds constraints that the running maximum has already

absorbed. The total cost is then the left-hand side of (4), so an unblocked feasible vector exists if and only if that cost does not exceed  $v_{\#}^0(n)$ . Any slack may be added to the largest payoff. ■

Now we are ready for our main existence theorem.

**Theorem 1** *Let  $\tau$  be an admissible threat map. Then  $ut\text{-}core^{\tau}(N, v)$  is non-empty for every symmetric superadditive game  $(N, v)$ .*

**Proof.** Normalize  $v(i) = 0$ , so that  $v \geq 0$  by superadditivity. By Lemma 3 it suffices to verify (4), which involves no threat map.

Let  $1 = m^1 < m^2 < \dots < m^p \leq n - 1$  be the *record sizes*, that is, the coalition sizes at which the running maximum  $\hat{c}$  strictly increases, and let  $d_t = c_{m^t} - c_{m^{t-1}} > 0$  for  $t \geq 2$  be the successive increments, so that  $\hat{c}_k = \sum_{t: m^t \leq k} d_t$ , the term  $t = 1$  being absent because  $c_{m^1} = c_1 = 0$ . Every increment is collected by  $\hat{c}_{n-1}$ , and  $d_t$  is collected by  $\hat{c}_k$  for exactly the  $n - m^t$  indices  $k$  with  $m^t \leq k \leq n - 1$ , so the left-hand side of (4) is

$$\hat{c}_{n-1} + \sum_{k=1}^{n-1} \hat{c}_k = \sum_{t \geq 2} d_t + \sum_{t \geq 2} d_t (n - m^t) = \sum_{t \geq 2} d_t (n - m^t + 1).$$

Now pack  $N$  greedily with disjoint blocks whose sizes are record sizes, largest first: fit as many disjoint blocks of size  $m^p$  as possible, then as many of size  $m^{p-1}$  into what remains, and so on down to  $m^2$ , leaving any remaining players as singletons. Let  $a_j$  denote the number of blocks of size  $m^j$ , and let  $A_t = \sum_{j \geq t} a_j m^j$  be the number of players covered by blocks of size at least  $m^t$ . Since at the stage corresponding to  $m^t$  the set of still uncovered players has fewer than  $m^t$  members, we have  $n - A_t \leq m^t - 1$ , that is,  $A_t \geq n - m^t + 1$  for every  $t \geq 2$ .

Because the blocks are disjoint and  $v \geq 0$ , superadditivity gives  $v_{\#}^0(n) \geq \sum_{j \geq 2} a_j v_{\#}^0(m^j) = \sum_{j \geq 2} a_j m^j c_{m^j}$ . Substituting  $c_{m^j} = \sum_{t=2}^j d_t$  and exchanging the order of summation,

$$v_{\#}^0(n) \geq \sum_{j \geq 2} a_j m^j c_{m^j} = \sum_{t \geq 2} d_t A_t \geq \sum_{t \geq 2} d_t (n - m^t + 1),$$

which is exactly the left-hand side of (4). ■

Since  $\mu$  is admissible, the ut-core of every symmetric superadditive game is non-empty; and since  $\rho$  is admissible too, so is the r-ut-core.

Anonymity can be dispensed with when the game is convex.

**Remark 1** *Let  $(N, v)$  be symmetric, superadditive and convex, and let  $\tau$  be covariant with  $\tau(S) \in \text{core}(S, v_S)$  for every  $S$  in its domain, but not necessarily anonymous. Then  $\text{ut-core}^\tau(N, v)$  is non-empty.*

To see this, normalize  $v(i) = 0$  and let  $m_k = v_\#(k) - v_\#(k-1)$ , which convexity makes non-decreasing. Under any core allocation of a  $k$ -player coalition the remaining  $k-1$  members receive at least  $v_\#(k-1)$  between them, so  $\tau_i(S) \leq m_k$  for every  $i \in S$ , and the coalition is defused once one member receives  $m_k$ . Imposing  $x_{(n-k+1)} \geq m_k$  for every  $k \leq n-1$  and repeating the computation in the proof of Lemma 3, the cheapest such vector costs  $m_{n-1} + \sum_{k=1}^{n-1} m_k$ , which telescopes to  $m_{n-1} + v_\#(n) - m_n \leq v_\#(n)$ .

Theorem 1 should be set against the corresponding statement for the core, which is immediate from Lemma 2.

**Corollary 1** *Let  $(N, v)$  be symmetric. Then  $\text{core}(N, v) \neq \emptyset$  if and only if  $c_n \geq \hat{c}_{n-1}$ , that is, if and only if per capita worth is maximal at the grand coalition.*

The symmetric case also settles a question about anonymity. Our solution is anonymous, but it does not select symmetric allocations. Under the equal split every player receives  $c_n$ , so by (5) it survives if and only if  $c_n \geq \hat{c}_{n-1}$ , which by Corollary 1 is exactly the condition for the core to be non-empty. As the equal split is the only symmetric vector distributing  $v(N)$ , whenever the core is empty and our solution is not, no element of the latter treats the players alike: the solution is symmetric as a set and every one of its elements breaks that symmetry. This is the formal content of the collusion story – somebody has to be excluded – and it is a feature, not a failure of anonymity.

The contrast has an important economic reading. Where the core additionally requires the last member of the grand coalition to contribute enough to keep per capita worth from falling, it otherwise predicts that his mere presence prevents an agreement,

because he can join blocking subcoalitions on arbitrary internal terms. Once threats must be credible and unconditional, agreement becomes feasible.<sup>14</sup>

Our next result shows the non-emptiness of the  $ut\text{-}core^\tau$  for three-player superadditive games, without any further restriction. This is in contrast to the well-known fact that superadditivity does not suffice for the core to be non-empty in three-player games – balancedness being what is needed (Shapley, 1967) – as the game  $v'$  of Example 2 illustrates.

**Theorem 2** *Let  $\tau$  be admissible and let  $(N, v)$  be superadditive with two or three players. Then  $ut\text{-}core^\tau(N, v)$  is non-empty.*

**Proof.** The two-player case is obvious, so let  $N = \{1, 2, 3\}$ . By covariance normalize  $v(i) = 0$ , so that  $v \geq 0$ , and relabel so that  $v(\{1, 2\}) \geq v(\{1, 3\}), v(\{2, 3\})$ . Only singletons and pairs can threaten, and admissibility pins both down: efficiency gives  $\tau(\{i\}) = 0$ , and Lemma 1 gives each member of a pair half that pair's worth.<sup>15</sup> Consider

$$x = \left( \frac{v(\{1, 2\})}{2}, \frac{v(\{1, 2\})}{2}, v(N) - v(\{1, 2\}) \right),$$

which is non-negative, since  $v(N) \geq v(\{1, 2\})$  by superadditivity, and distributes  $v(N)$ . Player 1 alone defeats the block by  $\{1, 2\}$  and by  $\{1, 3\}$ , as  $x_1 = v(\{1, 2\})/2 \geq v(\{1, 3\})/2$ , and Player 2 defeats the block by  $\{2, 3\}$  for the same reason. Hence  $x \in ut\text{-}core^\tau(N, v)$ . ■

Here every proper subcoalition is a singleton or a pair, so all admissible threat maps generate the same solution; in particular the recursive solution of Section 4.3 coincides there with the  $ut\text{-}core$ .

Our final existence result covers all the games in which the only coalitions that can block are pairs, but it is restricted to the  $ut\text{-}core$ . Three observations make this class wider than it may appear. First, a coalition outside the domain of  $\tau$  cannot block at

---

<sup>14</sup>When per capita worth is non-decreasing throughout, the core is non-empty as well, an interesting parallel with Moldovanu and Winter (1994), where convexity is a condition for non-cooperative implementation.

<sup>15</sup>For two players this coincides with the Nash Bargaining Solution.

all; for the core center these are exactly the coalitions with an empty core. Second, a coalition whose threat awards some member exactly her autarky payoff cannot block either. Third, efficiency makes large coalitions cheap to defuse, since some member's threat is at most  $v(S)/|S|$ .

**Lemma 4** *Let  $\tau$  be a threat map. If  $\tau(S)_i = v(i)$  for some  $i \in S$ , then  $S$  does not block any individually rational payoff vector.*

**Proof.** Blocking by  $S$  requires  $x_j < \tau(S)_j$  for every  $j \in S$ , and in particular  $x_i < v(i)$ , contradicting individual rationality. ■

**Theorem 3** *Let  $(N, v)$  be superadditive and suppose that every proper coalition  $S \subset N$  with  $|S| \geq 3$  either has an empty core or contains a player  $i$  who is a dummy in  $S$ , in the sense that  $v(S) = v(S \setminus i) + v(i)$ . Then  $ut\text{-}core(N, v) \neq \emptyset$ .*

**Proof.** We first check that no coalition of three or more can block. If  $core(S, v_S)$  is empty then  $S \notin \mathcal{D}(\mu)$ . Otherwise let  $i$  be a dummy in  $S$ ; every  $x \in core(S, v_S)$  satisfies  $x(S \setminus i) \geq v(S \setminus i)$  and  $x(S) = v(S)$ , so  $x_i \leq v(S) - v(S \setminus i) = v(i)$ , while individual rationality gives  $x_i \geq v(i)$ . Hence the core of  $S$  lies in the face  $x_i = v(i)$ , so  $\mu(S, v_S)_i = v(i)$  and Lemma 4 applies.

By covariance we may now normalize  $v(i) = 0$ , so that  $v \geq 0$  and, by Lemma 1, the threat of any pair  $\{i, j\}$  is  $(v(\{i, j\})/2, v(\{i, j\})/2)$ . Only pairs can block a non-negative payoff vector, so it suffices to exhibit a non-negative  $x$  with  $x(N) = v(N)$  such that every pair contains a member  $i$  with  $x_i \geq v(\{i, j\})/2$ .

Construct a greedy matching: let  $f_1 = \{a_1, b_1\}$  be a pair of maximal worth in  $N$ , let  $f_2 = \{a_2, b_2\}$  be a pair of maximal worth in  $N \setminus f_1$ , and so on, until fewer than two players remain. Order the players as  $a_1, b_1, a_2, b_2, \dots$ , placing any unmatched player last, and set

$$x_i = \max \{v(\{i, j\})/2 : j \text{ follows } i \text{ in this order}\},$$

with  $x_i = 0$  for the last player. Every pair is then covered by whichever of its two members comes first, so no pair blocks.

As for the budget, every player following  $a_k$  or  $b_k$  lies in  $N \setminus (f_1 \cup \dots \cup f_{k-1})$ , so every pair contributing to  $x_{a_k}$  or to  $x_{b_k}$  is a pair within that set, and hence has worth at most  $v(f_k)$  by the maximality of  $f_k$ . Therefore  $x_{a_k}, x_{b_k} \leq v(f_k)/2$ , and

$$\sum_{i \in N} x_i \leq \sum_k 2 \frac{v(f_k)}{2} = \sum_k v(f_k) \leq v(N),$$

the last inequality holding by superadditivity, since the  $f_k$  are disjoint. Any slack may be distributed arbitrarily. ■

Only the first step of the proof is specific to the core center. For any admissible threat map, the same construction works whenever no coalition of three or more can block; what the displayed hypothesis adds is that, for  $\mu$ , this can be checked on the characteristic function alone.

## 5 The three-player characterization

Three-player games are worth treating in full: the core and the ut-core routinely differ there, and both are simple enough to compute. Denote an arbitrary three-player, superadditive game by

$$v = (a, b, c; \alpha + a + b, \beta + a + c, \gamma + b + c; \alpha + a + b + c + \delta),$$

with all letters non-negative and, wlog,  $\alpha \geq \beta \geq \gamma$ . Note that the greek letters capture the marginal gain beyond additivity.

We can now characterize the solution completely. Since every proper subcoalition here has at most two members, all admissible threat maps generate the same solution, so the characterization is not specific to the core center. By covariance it suffices to treat the auxiliary game

$$\hat{v} = (0, 0, 0; \alpha, \beta, \gamma; \alpha + \delta),$$

obtained by subtracting  $a$ ,  $b$  and  $c$  from the payoffs of Players 1, 2 and 3, since  $\text{ut-core}(\{1, 2, 3\}, v) = (a, b, c) + \text{ut-core}(\{1, 2, 3\}, \hat{v})$ . Write  $\Delta = \{x \in \mathbb{R}_+^3 : x_1 + x_2 + x_3 = \alpha + \delta\}$  for the simplex of feasible payoffs.

**Proposition 4** *Let  $A = \{x_1 \geq \alpha/2\}$ ,  $A' = \{x_2 \geq \alpha/2\}$ ,  $B = \{x_1 \geq \beta/2\}$ ,  $B' = \{x_3 \geq \beta/2\}$ ,  $C = \{x_2 \geq \gamma/2\}$  and  $C' = \{x_3 \geq \gamma/2\}$ . Then*

- (i)  $\text{ut-core}(\{1, 2, 3\}, \hat{v}) = \Delta \cap [(A \cap C) \cup (A \cap C') \cup (A' \cap B) \cup (A' \cap B')]$ , a union of four convex polygons, each cut from  $\Delta$  by two lower bounds;*
- (ii) it is never empty, since it always contains  $(\alpha/2, \alpha/2, \delta)$ ;*
- (iii)  $\text{core}(\{1, 2, 3\}, \hat{v}) \neq \emptyset$  if and only if  $\beta + \gamma \leq \alpha + 2\delta$ .*

**Proof.** Only pairs and singletons can block. Singletons impose  $x \geq 0$ , and by Lemma 1 the threat of a pair is half its worth to each member, so  $x \in \Delta$  survives if and only if, for each pair, at least one member receives at least half that pair's worth:

$$\text{ut-core}(\{1, 2, 3\}, \hat{v}) = \Delta \cap (A \cup A') \cap (B \cup B') \cap (C \cup C').$$

Since  $\alpha \geq \beta$  we have  $A \subseteq B$ , and since  $\alpha \geq \gamma$  we have  $A' \subseteq C$ , so that

$$(A \cup A') \cap (B \cup B') \cap (C \cup C') = [A \cap (C \cup C')] \cup [A' \cap (B \cup B')],$$

each bracket absorbing the constraint the other branch has to carry. Distributing the two unions gives (i). For (ii),  $(\alpha/2, \alpha/2, \delta)$  lies in  $\Delta$  and in  $A \cap C$ , as  $\alpha/2 \geq \gamma/2$ . For (iii), rewrite the coalitional rationality constraints  $x_1 + x_2 \geq \alpha$ ,  $x_1 + x_3 \geq \beta$  and  $x_2 + x_3 \geq \gamma$  as the upper bounds  $x_3 \leq \delta$ ,  $x_2 \leq \alpha + \delta - \beta$  and  $x_1 \leq \alpha + \delta - \gamma$ , all non-negative because  $\alpha \geq \beta \geq \gamma$  and  $\delta \geq 0$ . A simplex of total  $\alpha + \delta$  meets such bounds if and only if they sum to at least that total, which gives  $\alpha + 2\delta \geq \beta + \gamma$ . ■

Part (iii) is worth setting beside (ii): the ut-core of a three-player superadditive game is never empty, while the core disappears as soon as the two smaller synergies together outweigh the largest one by more than  $2\delta$ . The geometry behind the four pieces is easiest to see by projecting  $\Delta$  onto the 1-2 plane, viewing Player 3's payoff as the vertical dimension.

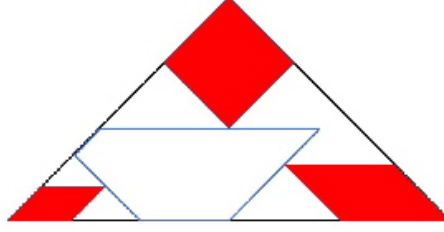


Figure 1

Figure 1 illustrates the game  $v = (0, 0, 0; 6, 5, 3; 10)$ . The white region is the ut-core, each parallelogram collecting the agreements removed by one pair's credible block, and the blue-bordered pentagon is the core. The parallelograms may overlap, and must do so when the core is empty,<sup>16</sup> but only in the interior: by superadditivity none extends beyond half an edge, so on the edges they can at most touch. Their own edges and vertices survive blocking, which is why the ut-core persists even when they cover the whole simplex, as in Example 2.

## 6 Examples

The examples below isolate the features of the solution that matter for economic analysis.

**Example 1** For  $v = (0, 0, 0; 1, 1, 0; 1 + \delta)$  with  $\delta \leq 1$ , the core is the blue parallelogram with the vertices  $(1 + \delta, 0, 0)$ ,  $(1, 0, \delta)$ ,  $(1 - \delta, \delta, \delta)$  and  $(1, \delta, 0)$ . See Figure 2, and  $\mu = (1, \delta/2, \delta/2)$ .<sup>17</sup>

<sup>16</sup>A simplex of total  $\alpha + \delta$  meets a box of strict upper bounds exactly when those bounds sum to more than the total, so the parallelograms of  $\{1, 2\}$  and  $\{1, 3\}$  meet if and only if  $2\beta > \alpha + 2\delta$ , and those of  $\{1, 2\}$  and  $\{2, 3\}$  if and only if  $2\gamma > \alpha + 2\delta$ . Were both to fail, adding them would give  $\beta + \gamma \leq \alpha + 2\delta$ , so the core would be non-empty by Proposition 4(iii).

<sup>17</sup>For  $\delta > 1$  the constraint  $x_1 \geq 0$  binds as well, and the core is the pentagon obtained by cutting the vertex  $(1 - \delta, \delta, \delta)$  off. Nothing below is affected: the ut-core is given by Proposition 4 for every  $\delta \geq 0$ .

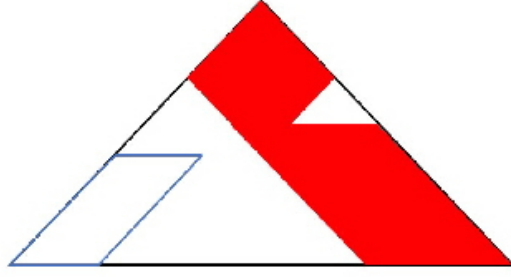


Figure 2

Here  $\alpha = \beta = 1$ ,  $\gamma = 0$ , so the pair threats are  $(1/2, 1/2)$  for  $\{1, 2\}$  and  $\{1, 3\}$  and  $(0, 0)$  for  $\{2, 3\}$ . Proposition 4 collapses to two components – overlapping when  $\delta > 1/2$  – in one of which Player 1 receives at least  $1/2$ , while in the other Players 2 and 3 each do:

$$\text{ut-core}(\{1, 2, 3\}, v) = \{x \in \Delta : x_1 \geq 1/2\} \cup \{x \in \Delta : x_2, x_3 \geq 1/2\}.$$

At  $\delta = 0$  the second component collapses to the single point  $(0, 1/2, 1/2)$ ; as  $\delta$  grows it opens up, and the two components overlap once  $\delta > 1/2$ .

What sustains the core here is that Player 1 may offer either partner a share of that partner's own payoff, on terms that depend on the deal she is attacking. Under unconditional threats she may only threaten to take one partner away and negotiate, which in expectation splits the surplus equally, so she can neither obtain nor offer more than  $1/2$ . The second component is therefore collusion between Players 2 and 3: by refusing to compete for Player 1 they extract surplus from her.

**Example 2** Let  $v' = (0, 0, 0; 1, 1, 1; 1)$ . We set  $\delta = 0$  for simplicity. Now the core is empty, so  $\mu$  does not exist for the game itself; for each pair, however,  $\mu = (1/2, 1/2)$ . Consequently, at least two players need to get at least a half to prevent credible blocking. There are three ways of choosing two players out of three, so our solution consists of three components (since  $\delta = 0$ , these are just points) depending on which player is left out:

$$\text{ut-core}(\{1, 2, 3\}, v') = (1/2, 1/2, 0) \cup (1/2, 0, 1/2) \cup (0, 1/2, 1/2).$$

Again, the player left out cannot entice one of the others to block the agreement, since he could not credibly offer more than  $1/2$ . Since we have no information to choose between the three components, it makes sense that all three are in principle possible.

**Example 3** Let  $v'' = (0, 0, 0; 1, 0, 0; 1)$ . Now the core is  $(x, 1 - x, 0)$ ,  $x \in [0, 1]$ , so  $\mu = (1/2, 1/2, 0)$ , the same as the Shapley value (Shapley, 1953). For the pairs,  $\mu(\{1, 2\}) = (1/2, 1/2)$  while  $\mu(\{1, 3\}) = \mu(\{2, 3\}) = (0, 0)$ , so

$$\begin{aligned} \text{ut-core}(\{1, 2, 3\}, v'') &= \{(x, y, 1 - x - y), \text{ with } x \geq 1/2, y \in [0, 1 - x]\} \cup \\ &\quad \{(x, y, 1 - x - y), \text{ with } y \geq 1/2, x \in [0, 1 - y]\} \end{aligned}$$

the two white triangles.

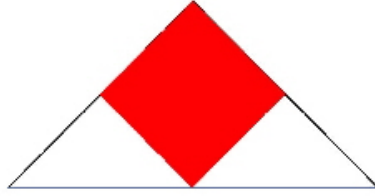


Figure 3

Note that 3 is a dummy player – his presence changes the value of no coalition he joins – yet he has a positive payoff. Normatively one might insist that a dummy earn his autarky value; positively it may be wrong to ignore his presence. As long as either 1 or 2 earns at least  $1/2$ , the pair cannot block, so it is as if 3 had bribed one of them.

It is instructive to set the three examples beside the standard alternatives. The  $\varepsilon$ -core is reported at its least value, and Example 1 is evaluated at  $\delta = 0$ .

|               | Example 1 ( $\delta = 0$ )                                        | Example 2                                  | Example 3                                                |
|---------------|-------------------------------------------------------------------|--------------------------------------------|----------------------------------------------------------|
| core          | $(1, 0, 0)$                                                       | empty                                      | $\{(x, 1 - x, 0)\}$                                      |
| least core    | $(1, 0, 0)$                                                       | $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  | $\{(x, 1 - x, 0)\}$                                      |
| nucleolus     | $(1, 0, 0)$                                                       | $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  | $(\frac{1}{2}, \frac{1}{2}, 0)$                          |
| Shapley value | $(\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$                         | $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  | $(\frac{1}{2}, \frac{1}{2}, 0)$                          |
| ut-core       | $\{x_1 \geq \frac{1}{2}\} \cup \{(0, \frac{1}{2}, \frac{1}{2})\}$ | $(\frac{1}{2}, \frac{1}{2}, 0)$ and perms. | $\{x_1 \geq \frac{1}{2}\} \cup \{x_2 \geq \frac{1}{2}\}$ |

Two contrasts stand out. In Example 1 the core, the least core and the nucleolus all reduce to the single point  $(1, 0, 0)$ : Player 1 takes everything, the competition between Players 2 and 3 being resolved entirely in her favour. Ours is the only solution containing agreements in which they do not compete, and it identifies those agreements as a component in their own right rather than as a compromise. In Example 3 every alternative awards the dummy nothing, whereas ours pays him whenever one of the other two is paid at least a half.

Example 3 sits in a one-parameter family that we can use to compare the two threat maps. For  $w^\delta = (0, 0, 0; 1, 0, 0; 1 + \delta)$  with  $\delta \geq 0$ , the core is  $\{x \in \Delta : x_3 \leq \delta\}$  and, by Proposition 4, the ut-core is  $\Delta \setminus \{x_1 < 1/2, x_2 < 1/2\}$ , the simplex of total  $1 + \delta$  with a corner square removed. Integrating over each,

$$\mu_1(w^\delta) = \frac{(1 + \delta)^3 - 1}{3[(1 + \delta)^2 - 1]} \quad \text{and} \quad \rho_1(w^\delta) = \frac{8(1 + \delta)^3 - 3}{24(1 + \delta)^2 - 12},$$

with  $\mu_2 = \mu_1$ ,  $\rho_2 = \rho_1$  and the third coordinate given by (1); for  $\rho$  we have used Proposition 3(iii), by which the r-ut-core of a three-player game is its ut-core. At  $\delta = 0$  the core is the segment  $x_3 = 0$  and the first expression is to be read on that segment, giving  $\mu = (1/2, 1/2, 0)$  and  $\rho = (5/12, 5/12, 1/6)$ , as computed in Example 3. At  $\delta = 1$ ,

$$\mu(w^1) = \left(\frac{7}{9}, \frac{7}{9}, \frac{4}{9}\right) \quad \text{and} \quad \rho(w^1) = \left(\frac{61}{84}, \frac{61}{84}, \frac{23}{42}\right).$$

In both cases the recursion shifts weight towards Player 3:  $\rho_3 > \mu_3$  while  $\rho_1 < \mu_1$ . Embedding two members of this family in a single four-player game turns that crossing into a crossing of solutions.

**Example 4** Let  $N = \{1, 2, 3, 4\}$  with  $v(i) = 0$  for every  $i$ ,

$$v(\{1, 2\}) = 1, \quad v(S) = 0 \text{ for every other pair}, \quad v(\{1, 2, 3\}) = 2, \quad v(\{1, 2, 4\}) = 1,$$

$v(\{1, 3, 4\}) = v(\{2, 3, 4\}) = 0$  and  $v(N) = 2$ . The game is superadditive. Restricted to  $\{1, 2, 3\}$  it is  $w^1$ , and restricted to  $\{1, 2, 4\}$  it is  $w^0$ , the game of Example 3; the remaining two triples are worthless, as are all pairs other than  $\{1, 2\}$ , whose threat is  $(1/2, 1/2)$  under either map.

Only  $\{1, 2\}$  and  $\{1, 2, 3\}$  bind. The triple  $\{1, 2, 4\}$  is harmless under  $\mu$  by Lemma 4, since Player 4 is a dummy in it, and its threat under  $\rho$ ,  $(5/12, 5/12, 1/6)$ , is too small to bite at either of the vectors below; the worthless coalitions threaten with the zero vector under both maps. Consider

$$x = \left(\frac{7}{10}, \frac{7}{10}, \frac{1}{2}, \frac{1}{10}\right) \quad \text{and} \quad y = \left(\frac{3}{4}, \frac{3}{4}, \frac{2}{5}, \frac{1}{10}\right),$$

both of which distribute  $v(N) = 2$  and neither of which is blocked by  $\{1, 2\}$ . For  $x$ , the triple  $\{1, 2, 3\}$  does not block under  $\mu$ , because  $x_3 = 1/2 \geq 4/9$ , but it does block under  $\rho$ , because  $7/10 < 61/84$  and  $1/2 < 23/42$ . For  $y$  the comparison is reversed:  $\{1, 2, 3\}$  blocks under  $\mu$ , as  $3/4 < 7/9$  and  $2/5 < 4/9$ , but not under  $\rho$ , as  $3/4 \geq 61/84$ . Hence

$$x \in \text{ut-core}(N, v) \setminus \text{r-ut-core}(N, v) \quad \text{and} \quad y \in \text{r-ut-core}(N, v) \setminus \text{ut-core}(N, v).$$

**Proposition 5** *There are superadditive games for which neither  $\text{r-ut-core}(N, v) \subseteq \text{ut-core}(N, v)$  nor  $\text{ut-core}(N, v) \subseteq \text{r-ut-core}(N, v)$ .*

**Proof.** Example 4. ■

The collusion of Example 1 is not confined to three players, and the scope for it turns on the technology rather than on the size of the game.

**Example 5** Let  $N = \{1, 2, 3, 4\}$  and let  $v(S) = 1$  whenever  $S$  contains Player 1 together with at least two of the others, and  $v(S) = 0$  otherwise. The game is superadditive, since no two disjoint coalitions can both contain Player 1.

Its core is the single point  $(1, 0, 0, 0)$ : for each pair  $\{j, k\} \subset \{2, 3, 4\}$  the constraint  $x(\{1, j, k\}) \geq 1$  drives the remaining player's payoff to zero. The collapse is a four-player phenomenon. Restricted to any one triple  $\{1, j, k\}$  the game is worth 1 while every pair inside it is worth nothing, so that subgame's core – and its ut-core, since nothing there can block – is the entire simplex. It is the three triples together that leave Player 1 with everything.

Those same triples are the only coalitions that can block in the grand coalition. Every pair is worthless, as is  $\{2, 3, 4\}$ , and each  $\{1, j, k\}$  is symmetric once normalized

– each member alone and each pair within it are worth nothing, the three together are worth 1 – so Lemma 1 fixes its threat at  $(1/3, 1/3, 1/3)$ , whatever admissible threat map is used. Hence

$$\text{ut-core}^\tau(N, v) = \{x \in \Delta : x_1 \geq \tfrac{1}{3}\} \cup \bigcup_{j < k} \{x \in \Delta : x_j, x_k \geq \tfrac{1}{3}\},$$

the union running over the three pairs from  $\{2, 3, 4\}$ , and every component is full-dimensional. Here too the structure is irreducibly four-player: a triple threatens only as a *proper* subcoalition, and what bites is the overlap of three of them.

Read Player 1 as a firm with a job requiring two of the three workers 2, 3, 4, and the two halves of the solution are the two readings of that situation. The core records complete competition: the workers undercut one another and the firm takes the entire surplus. The second family of components records collusion: any two workers who each hold out for a third cannot be separated, since all the firm can threaten is to walk into a subgame whose surplus it expects to split three ways. Two workers suffice, the third being free to be paid nothing, so the cartel need not be comprehensive.

That collusion is sustained by the technology rather than by the number of players. Had the job required only one worker, each triple  $\{1, j, k\}$  would reduce to the game of Example 1 with  $\delta = 0$ , whose threat awards the two workers nothing and is therefore harmless by Lemma 4, leaving only the pairs  $\{1, j\}$  to threaten with  $(1/2, 1/2)$ . Collusion would then require all three workers to hold out for a half, at a cost of  $3/2$ , which is infeasible; the ut-core would collapse to  $\{x_1 \geq 1/2\}$  and the firm’s position would be restored. What sustains collusion is not how many workers there are but how many of them the firm must engage.

## 7 Concluding remarks

We have presented a modified version of the classical solution concept of the core, based on requiring blocking threats to be unconditional, representing the expected outcome of the blocking coalition “going it alone”. The requirement turns any solution concept into another, so what we obtain is a family rather than a single object, every member

of which contains the core and is non-empty in situations where the core is not.

The softening of competition that this produces gives rise to phenomena the core cannot express. Players with little to contribute may take a disproportionate share when the others are already earning enough not to block, and players who would otherwise be played off against one another can gain by declining to compete. Example 5 makes this precise: where the core hands a firm the whole surplus because three workers undercut each other, any two of them who hold out cannot be separated. Collusion is a real phenomenon, and a solution concept that can represent it is the better for it.

Asking a coalition to expect from its subgame what we predict for the whole game turns the requirement on the construction itself. That fixed point exists and is unique, and it neither refines nor extends the solution generated by the core, so the choice of what a departing coalition expects is not a formality.

Two questions we leave open. The first is non-emptiness: we have shown that it holds for symmetric superadditive games, for superadditive games with at most three players, and when only pairs can block, but whether some superadditive game with four or more players has an empty ut-core we do not know.

The second is the extension beyond transferable utility. Nothing in the construction is specific to it: replacing  $v$  by a coalitional feasible-set correspondence leaves every definition intact, and the inclusion of the core is immediate. What does not survive is efficiency. A solution of a TU subgame lies in a hyperplane, so the mean of any belief over it distributes  $v(S)$  exactly; a non-transferable frontier is curved, so the mean lies strictly inside the feasible set when that set is convex, and outside it when it is not. Since efficiency of threats is what drives Lemma 1 and every existence result that follows, an extension to games without side payments needs a different engine.

## References

- [1] Aumann, Robert J. “The core of a cooperative game without side payments.” *Transactions of the American Mathematical Society* 98 (1961): 539-552.
- [2] Aumann, Robert J., and Michael Maschler. The bargaining set for cooperative games. *Advances in game theory* Vol. 10. Princeton, NJ: Princeton University, (1964): 443-476.
- [3] Bennett, Elaine. “The aspiration approach to predicting coalition formation and payoff distribution in sidepayment games.” *International Journal of Game Theory* 12 (1983): 1-28.
- [4] Bennett, Elaine, and William R. Zame. “Bargaining in cooperative games.” *International Journal of Game Theory* 17 (1988): 279-300.
- [5] Burguet, Roberto, and Ramon Caminal. “Coalitional bargaining with consistent counterfactuals.” *Journal of Economic Theory* 187 (2020): 105013.
- [6] Chwe, Michael Suk-Young. “Farsighted coalitional stability.” *Journal of Economic Theory* 63(2) (1994): 299-325.
- [7] Cross, John G. “Some theoretic characteristics of economic and political coalitions.” *Journal of Conflict Resolution* 11 (1967): 184-195.
- [8] Dutta, Bhaskar, Debraj Ray, Kunal Sengupta, and Rajiv Vohra. “A consistent bargaining set.” *Journal of Economic Theory* 49(1) (1989): 93-112.
- [9] Edgeworth, Francis Ysidro. *Mathematical psychics: An essay on the application of mathematics to the moral sciences*. Kegan Paul, London (1881).
- [10] Gillies, Donald B. “Some theorems on n-person games.” *Doctoral Thesis* (1953). Department of Mathematics, Princeton.
- [11] Gillies, Donald B. “Solutions to general non-zero-sum games.” *Contributions to the Theory of Games* 4, no. 40 (1959): 47-85.

- [12] González-Díaz, Julio and Estela Sánchez-Rodríguez. “A natural selection from the core of a TU game: The core-center.” *International Journal of Game Theory* 36 (2007): 27-46.
- [13] González-Díaz, Julio and Estela Sánchez-Rodríguez. “Towards an axiomatization of the core-center.” *European Journal of Operational Research* 195(2) (2009): 449-459.
- [14] Harsanyi, John C., and Reinhard Selten. *A general theory of equilibrium selection in games*. MIT Press Books, Cambridge MA (1988).
- [15] Kalai, Ehud, Andrew Postlewaite, and John Roberts. “A Group Incentive Compatible Mechanism Yielding Core Allocation.” *Journal of Economic Theory* 20 (1979): 13-22.
- [16] Maschler, Michael, Bezalel Peleg, and Lloyd S. Shapley. “Geometric properties of the kernel, nucleolus, and related solution concepts.” *Mathematics of Operations Research* 4 (1979): 303-338.
- [17] Moldovanu, Benny, and Eyal Winter. “Core implementation and increasing returns to scale for cooperation.” *Journal of Mathematical Economics* 23(6) (1994): 533-548.
- [18] Okada, Akira, and Eyal Winter. “A Non-cooperative Axiomatization of the Core.” *Theory and Decision* 53(1) (2002): 1-28.
- [19] Peleg, Bezalel. “Axiomatizations of the core.” in, *Handbook of game theory with economic applications* 1. (R.J. Aumann and S. Hart eds.) North-Holland. (1992): 397-412.
- [20] Platt, Brennan C. “Spoilers, blocking coalitions, and the core.” *Social Choice and Welfare* 33(3) (2009): 361-381.
- [21] Ray, Debraj. “Credible coalitions and the core.” *International Journal of Game Theory* 18 (1989): 185-187.

- [22] Ray, Debraj, and Rajiv Vohra. "Equilibrium binding agreements." *Journal of Economic Theory* 73(1) (1997): 30-78.
- [23] Schmeidler, David. "The nucleolus of a characteristic function game." *SIAM Journal on Applied Mathematics* 17(6) (1969): 1163-1170.
- [24] Selten, Reinhard. "A noncooperative model of characteristic-function bargaining." In: *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern* (1981): 131-151.
- [25] Shapley, Lloyd S. "A value for  $n$ -person games." in *Contributions to the Theory of Games II* (H. W. Kuhn and A. W. Tucker eds.), *Annals of Mathematics Studies* 28, Princeton University Press (1953): 307-317.
- [26] Shapley, Lloyd S., and Martin Shubik. "Quasi-cores in a monetary economy with nonconvex preferences." *Econometrica* 34 (1966): 805-827.
- [27] Shapley, Lloyd S. "On balanced sets and cores." *Naval Research Logistics Quarterly* 14(4) (1967): 453-460.