Consistent credible blocking and the core*

József Sákovics

Universitat de les Illes Balears

April 20, 2026

Abstract

The core is one of the most widely used cooperative solution concepts in applications to the social sciences, providing a rigorous formulation of Edgeworth's original insight on the stability of surplus allocation. By requiring that no coalition can "block" a proposed allocation, the standard definition of the core implicitly allows blocking coalitions to coordinate on arbitrary distributions of their joint payoff, tailored to the allocation under consideration. We question whether this degree of flexibility adequately reflects the notion of a credible blocking agreement. In response, we require that admissible blocking agreements be consistent, in the sense that they coincide with the expected outcome of the subgame that arises when the members of a blocking coalition leave the grand coalition. Naturally, these outcomes are independent of the allocation being challenged. The resulting solution concept is a superset of the core, providing a solution in practically relevant applications where the core is empty. Specifically, we show that it is non-empty for all symmetric, superadditive transferable utility

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

games satisfying a convexity condition, as well as for all three-player superadditive games. From an economic perspective, our solution captures the possibility of collusion among players in a competitive scenario, and the influence of non-contributing players on the outcome.

1 Introduction

A central feature of cooperative games is that players can form coalitions and enter into binding agreements on how to distribute the surplus generated. In view of this description, we reinterpret the definition of one of the most widely used solution concepts in applications of cooperative game theory: the core (Gillies, 1953, 1959). Our approach is explicitly normative. Our innovation is to impose a *consistency* requirement on admissible "blocking threats", given that they are to be interpreted as binding agreements reached by the blocking coalition. We also outline a dynamic motivation, in which such threats must be *credible* in a sequentially rational sense, but without leaving the sphere of cooperative analysis.

The core is a solution for cooperative games, which nonetheless was not originally defined axiomatically – although axiomatizations have subsequently been provided (see, for example, Peleg, 1992). Rather, it reflects a natural stability requirement, first pointed out by Francis Edgeworth (1881): for allocations agreed upon by the grand coalition no subcoalition should be able to secure an alternative agreement that makes all its members strictly better off.

The precise meaning of this requirement deserves closer scrutiny. What does it mean for a subcoalition to "secure" an alternative agreement? The standard interpretation holds that a coalition can block a proposed allocation whenever its worth exceeds the total payoff putatively assigned to its members. We argue that this interpretation is not the only one compatible with the underlying idea of binding agreements, emphasized at the outset.

We therefore propose a new solution concept based on an alternative understanding of what binding agreements entail in this setting. This solution captures the notion of consistency across threats made by the same coalition. While our objective is not merely to extend the core, we also show that our solution exists for a strictly larger class of transferable-utility games than those for which the core is nonempty. This allows us to obtain meaningful predictions in economically relevant environments where the core itself fails to provide a solution.

The interpretive issue we raise regarding "binding agreements" is best illustrated by a simple example. Suppose Ann, Bob, and Cai are negotiating a joint agreement. In the spirit of the core, they consider candidate allocations and examine whether any subcoalition would wish to block them. A coalition blocks if its members can secure a binding agreement that makes them all better off.

Our point of departure is the observation that such side agreements should depend only on what the coalition can achieve on its own, and not on the particular allocation currently under consideration. This is *not* how the core is typically applied. To see this, consider three players and a "two-person job" in which individual players generate no surplus, any pair generates 2, and the grand coalition also generates 2. The core is empty, since no distribution of 2 among three players can give every pair at least 2.

But note that this conclusion relies on allowing blocking coalitions to tailor their agreements to the allocation they seek to block. For instance, suppose a putative allocation assigns 0 to Ann and 1 to both Bob and Cai. The coalition of Ann and Bob can only block by agreeing on a division that gives Ann $1 - \varepsilon$ and Bob $1 + \varepsilon$, for some $\varepsilon \in (0, 1)$. By contrast, if the putative allocation assigns 1 to both Ann and Cai and 0 to Bob, then Ann and Bob must instead agree on a division that gives Ann $1 + \varepsilon$ and Bob $1 - \varepsilon$. Thus, the same coalition relies on different "binding agreements" depending on the allocation under consideration. We claim that this is difficult to reconcile with the idea that agreements are determined by the coalition's own opportunities. A coalition's threat should be consistent across situations.

A complementary argument, closer in spirit to non-cooperative reasoning, but still short of a non-cooperative analysis, reinforces this concern. Especially thinking about a negotiation in large groups, with many possible blocking coalitions, the idea that subcoalitions can form and reach a commitment to a side deal *during* the negotiation in the grand coalition seems suspect. A more natural assumption, therefore, is that subcoalitions must first leave the grand coalition before forming internal agreements. Then they cannot credibly threaten with arbitrary side agreements, since such agreements may be renegotiated once the split occurs. Their threat is thus simply to exit, and it is blocking only if the expected outcome of the ensuing subgame makes all members better off.

It is important to underline that while in our second interpretation we appeal to sequential rationality, we are not doing so in the context of some underlying non-cooperative game. To see this, it is instructive to compare our approach with subgame perfection in non-cooperative equilibrium analysis. Subgame perfection eliminates threats of actions that are not sequentially rational but whose presence sustains some equilibrium outcomes, and as a result it *refines* the set of Nash equilibria. We, on the other hand, eliminate non-credible threats that block non-core outcomes, and as a result we *enlarge* the core set.

To formalize our take on consistent threats, we develop our solution in two steps, so that the relevance of each assumption is clear. First, we restrict blocking threats to allocations that belong to the core of the corresponding subgame, without yet imposing consistency across deviations. We show that this restriction leaves the core unchanged, generalizing the invariance result of Ray (1989).

Second, we impose consistency by requiring that each coalition rely on a fixed element of its subgame core when formulating blocking threats. To select this element, we adopt the interpretation outlined above and use the center of gravity of the core set,[1] that is, the expected allocation under a uniform[2] prior. We refer to the resulting solution concept as the ccb-core (core with *c*onsistently *c*redible *b*locking). We establish two existence results. The ccb-core is non-empty for all superadditive three-player games and for all symmetric superadditive games satisfying an auxiliary condition on economies of scale.

We complete the paper with a characterization of our solutions for three-player games – for which they actually coincide. The qualitative description is followed up with a sequence of examples that highlight the novel characteristics of the solution, relevant for economic analysis.

2 Relation to the literature

There are two strands of literature to which our contribution is related, although, strictly speaking, it belongs only to the one we discuss first. The first strand concerns the analysis of restrictions on outcomes that may be used to block an agreement in a cooperative game. We discuss here the contributions most closely connected to our approach.

In view of the possible emptiness of the core, Aumann and Maschler (1964) introduced the idea that when a blocking proposal – termed an objection – is put forward, a player excluded from the blocking coalition may respond with a counter-objection, thereby neutralizing the original objection. The resulting solution is the bargaining set,[3] which is non-empty.[4] Importantly, in this framework it is the players excluded

[1]The center of gravity of the core, called the core center, was first proposed and analyzed in (González-Díaz and Sánchez-Rodríguez, 2007).

[2]In the absence of further information, we have no basis to assign different probabilities to different elements of the solution set. Nonetheless, our solution concept could be easily modified for alternative, possibly behavioral, expectation functions.

[3]Strictly speaking, it is the pre-bargaining set, as individual rationality has to be imposed separately.

[4]There are a number of further solutions that have evolved from this idea, like the *kernel* or the

from a coalition who influence outcomes, a feature that is closer in spirit to coalition-formation bargaining models. By contrast, just as in the case of the core, we restrict attention to the behavior of players within the blocking coalition itself.

Dutta et al. (1989) extend this idea by allowing counter-objections to be recursively challenged. Their "consistent" bargaining set applies the same test to each objection in a chain, mirroring the recursive structure of our rccb-core. This solution lies between the core and the bargaining set – it contains the former and is contained in the latter – and it may be empty.

Ray (1989) is the closest precursor of our study. He also studies the credibility of blocking coalitions, as we do, but adopts a qualitatively different approach. A coalition – rather than a side agreement as in our case – is deemed credible if its members can reach an agreement that is itself immune to blocking by any credible subcoalition. Remarkably, he shows that restricting attention to such credible coalitions leaves the core unchanged. In our view, Ray's credibility requirement is insufficient, for two, related, reasons that stem from the set-valued nature of the core. First, establishing credibility requires more than the existence of some agreement within the blocking coalition; it requires that the coalition be able to implement the specific agreement used to block a given imputation. As we show, even this apparently stronger condition would not alter the solution. Second, and more importantly, we believe that a set-valued solution should be interpreted as reflecting indeterminacy rather than discretion: the model does not specify which element will be selected, rather than allowing for an arbitrary choice among them. Players can therefore only form beliefs over the elements of the solution set when anticipating "future play".

Burguet and Caminal (2020) take another cooperative solution concept, the Nash Bargaining Solution (NBS), as their point of departure, focusing on the determination of disagreement points. They argue that it is incoherent for different players to base their threats on distinct counterfactual outcomes. Their solution concept – So-

nucleolus but their discussion here would make us lose the focus of this review.

lution with Consistent Outside Options (SCOOP) – imposes consistency by requiring that all players share a common expectation about the outcome in the event of disagreement within any coalition. The resulting solution is probabilistic, with multiple coalitions forming with positive probability.[5] Somewhat surprisingly, SCOOP is generically unique in superadditive games. Although the NBS differs conceptually from the core, the underlying requirement that threats be credible and consistent is closely related to our own.

The second strand of literature concerns non-cooperative implementations of the core in the spirit of the Nash program, pioneered by Kalai et al. (1979). Our paper does not belong to this tradition, for two reasons. The first one is methodological: we do not construct a non-cooperative game and analyze its equilibria, rather we reinterpret the concept of a blocking threat within the cooperative framework. Second, we change the direction of causality: instead of looking for a scaffold to support the notion of the core, as the implementation literature does, we start with proposing a reinterpretation of how best to capture the original insight of Edgeworth (1881) on unstable allocations, and then analyze its consequences, even if it leads us away from the core. Nonetheless, our paper is closely related to this literature, as the latter has explored similar concerns to ours, like sequential rationality or stationarity of equilibria. Platt (2009) provides a useful overview.

Finally, due to the methodology they employ, Okada and Winter (2002) deserve special mention. Rather than specifying a particular game, they characterize a class of non-cooperative games through axioms that are sufficient to implement the core.[6] Notably, one of their key axioms is subgame consistency, which requires identical behavior across isomorphic subgames. This condition, first introduced by Harsanyi and Selten (1988), appears closely related to our requirement that a coalition select the same agreement regardless of the outcome under consideration for blocking. Neverthe-

[5] While sometimes the SCOOP is deterministic, when the core is empty it is necessarily probabilistic.

[6] In this sense, their paper is sort of in between the above two literatures, though clearly closer to the latter.

less, it leads to a different solution concept. This contrast highlights that seemingly similar restrictions may have different implications in cooperative and non-cooperative settings, largely because the former excludes certain agreements, whereas the latter sustains them.

3 Definitions

We denote a cooperative game with transferable utility by (N, v), where N is the set of players indexed by $1, 2, ..., n$, and the *characteristic function*, $v(.) : 2^N \setminus \emptyset \to \mathbb{R}$, 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.[7]

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 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 *assignation* 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* of the game is a (possibly empty) set of payoff vectors for all the players. The *core* of (N, v) is a solution 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

[7] 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).

core is "empty"):

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

The *core center* of (N, v), denoted by $\mu(N, v)$, assigns to each player their expected utility calculated over their payoffs in 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 $i \in S \subseteq N$.

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

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

4 The solution with consistent credible blocking

We develop our proposed solution in two steps. First, we require that any threat used to block an assignment in the grand coalition be credible, meaning that it lies in the core of the subgame played by the blocking coalition. As we discuss below, this requirement is, in principle, stronger than Debraj Ray's (1989) modified core, which is also designed to capture credibility. Nonetheless, we show that the restriction is ultimately nonbinding: the set of outcomes supported by credible blocking coincides with the core.

In the second step, we impose consistency. Specifically, each coalition must employ the same threat – still selected from the core of the subgame – whenever it attempts to block. We propose using the core center of the subgame as this common threat. This approach typically yields a superset of the core and has the advantages of being both tractable and easy to interpret.

4.1 Credible blocking

Our notion of credibility requires that any side agreement used to block the putative solution should belong to the solution of the subgame where the blocking coalition is on its own.[8] As our starting point for the set of possible solutions to any game is the core, the most straightforward way to achieve this is to ensure that the threat itself belongs to the core of the subgame following the splitting away of the putative blocking coalition.[9] 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* $x_T \in$ *core*(T, v_T)*.*

Also concerned about the credibility of blocking, Ray (1989) uses a different approach when defining his *modified core.* He qualifies coalitions, rather than coalitional agreements, as credible or not. He starts with defining singletons as credible coalitions and he calls a larger coalition credible if and only if there is a way to distribute its worth in a way that is immune to blocking by credible coalitions. This approach does not require that any blocking side-agreement be in the modified core, only that the modified core of the subgame is non-empty. Therefore, if we remove the recursivity

[8] This reflects what can be credibly sustained as a threat in the absence of commitment power at the moment of blocking. At the same time, credibility does not imply that blocking will occur, since the threat is determined independently of any *status quo*. Consequently, our notion of credibility does not necessarily impose a binding constraint on the solution. 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.

[9] 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.

– for the moment – from his definition, we are left with the requirement that only coalitions with a non-empty core can block. Since removing the recursion weakly enlarges the potential blocking coalitions, but still restricts them relative to the original core, the resulting solution must be a superset of the core and a subset of the modified core. As Ray (1989) shows that the core and the modified core actually coincide, the non-recursive modified core must also coincide with the core. We now strengthen[10] the requirement that the blocking coalition's core is non-empty to include the credibility restriction on the blocking threats themselves, and show that it still does not enlarge the core.

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

Proof. It is obvious that restricting the potential blocking threats can only (weakly) enlarge the solution set. To show that it does not strictly enlarge it, suppose that $x \in \mathbb{R}^N$ is not in $\text{core}(N, v)$. We will show that then there must exist $T \subset N$ and $y_T \in \mathbb{R}^T$, such that $y_T \in \text{core}(T, v_T)$ and y_T strictly dominates x_T. Start by noting that $x \notin \text{core}(N, v)$ implies that there exists $S \subset N$ such that $x(S) < v(S)$. Given such an S, take $T \subseteq S$, such that $x(T) < v(T)$, but x_T satisfies the "CR condition": for all $W \subset T$, $x(W) \geq v(W)$. To see that such a T always exists, start with the singletons. Since they have no strict subsets, they trivially satisfy the "CR condition", so if x is not IR, the proof is complete. Otherwise, the "CR condition" is satisfied for all doubletons, so either we have found a doubleton T that blocks, or the "CR condition" is satisfied for all triples, etc. Since $x(S) < v(S)$, at some point we must satisfy both conditions.

Now modify the payoff vector x_T by adding some positive terms, $z_i > 0$ for $i \in T$, so that $x(T) + \sum_{i \in T} z_i = v(T)$. This new payoff vector, call it y_T, satisfies $y_T(T) = v(T)$ and $y_T(W) \geq v(W)$ for all $W \subset T$, thus $y_T \in \text{core}(T, v_T)$. Moreover, by construction, y_T strictly dominates x_T, so the claim is proven. ■

[10] In principle, it could happen that a putative assignment is only blocked by a threat outside of the (non-empty) core.

While, for greater transparency, we have stated the above invariance result in its simplest form, it is easy to see that using the credibility restriction recursively would not affect the solution set either. Recall that in the proof we constructed a blocking coalitional agreement (y_T) that belongs to the core of the subgame for the blocking coalition T. This blocking agreement must also belong to the *recursive* core of the subgame, which is immune only to *credible* blocking, since – as above – restricting the admissible blocking threats can only enlarge to solution set of the subgame.[11]

The apparent invariance of the solution to restricting blocking agreements to themselves not being blocked may suggest that the core 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* 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. Naturally, this object should not be contingent on the assignment it is intended to block. The remainder of the paper explores the implications of this consistency requirement.

4.2 The consistent credible solution

We now impose that – in addition to credibility – any threat invoked by a subcoalition S to block a proposed agreement in the grand coalition coincide with the expected outcome of the subgame played by S. As this outcome only depends on (S, v_S), it implicitly ensures consistency across blocking threats (to different putative agreements).

In specific applications there may be reasons to believe that some outcomes within the solution set (the core) are more likely than others. In general, however, neither the analyst nor the players possess a well-defined prior over core allocations. A natural

[11] In fact, a recursive application of the proof above shows that it does not enlarge it anyway, but it is not needed for the demonstration.

benchmark, therefore, is to assume a uniform prior, which implies that the expected outcome coincides with the core's center of gravity – also known as the core center (González-Díaz and Sánchez-Rodríguez, 2007).

Definition 2 *A blocking threat $x_T \in \mathbb{R}^T$ by coalition T is* consistent credible *if and only if $x_T = \mu(T, v_T)$.*

On this basis, we define our consistent credible solution as follows:

Definition 3 *A payoff vector, $x \in \mathbb{R}^N$, is an element of* ccb-core(N, v) *if and only if it distributes the value of the grand coalition, and there is no core-center of some subcoalition that strictly dominates it*[12] *for the members of that subcoalition:*

$$\textit{ccb-core}\,(N,v) = \left\{x \in \mathbb{R}^N : x(N) = v(N) \text{ and } x_S \not< \mu\,(S, v_S) \text{ for all } S \subset N\right\}.$$

If a single-valued solution is preferred for practical purposes, then in line with the above definition, one can compute the center of gravity of the ccb-core, thereby obtaining the *ccb-core center*.

Since the core is well-known to be convex, its center of gravity is contained in it. Thus, in the ccb-core we are selecting from the sets of agreements (cores) from which the credible threats of blocking coalitions can be chosen. Then, it is immediate that the ccb-core is a superset of the core. To see that the relation may be strict, consider the following three-player game: $v(0, 0, 0; 8, 8, 8; 8)$. This game has an empty core, as in order to ensure that each two-player coalition earns at least 8, we would need at least 12 units for the grand coalition: $(4, 4, 4)$. At the same time, $(0, 4, 4)$, $(4, 0, 4)$, and $(4, 4, 0)$ are (the only) elements of ccb-core, since the credible threat by any potential blocking coalition is $\mu\left(\{i, j\}, v(0, 0; 8)\right) = (4, 4)$, so it is sufficient to ensure that no pair of players earn strictly less than 4 each. $(0, 4, 4)$ and its permutations are not

[12] We use strict domination, as in our set-up of a point-valued threat the players who make a strict gain cannot compensate those who are indifferent.

in the core because any two players can threaten to agree on $(4-\varepsilon, 4+\varepsilon)$, for some $\varepsilon \in (0,4)$. However, this is not credible, since we (and the players) expect that the subgame would lead to an agreement of $(4,4)$.

From the example it is easy to see that the threat points generated by non-singleton subcoalitions work differently from unilateral threats. The constraint that single player "coalitions" do not block an assignment – referred to as Individual Rationality – imposes that solutions be in the "quadrant" of the utility space such that each coordinate is above the autarky point, where players earn what they can generate on their own. In the case of consistent credible blocking, the relation is flipped: We only rule out the solutions that are in the quadrant such that each coordinate is lower than the threat point. This happens because to lose credibility it is sufficient that a single member of the subcoalition does not prefer to block, so for consistent credible blocking we need all of them to be willing to do so. Of course, there is also the additional complication that there are a number of threat points – as opposed to the single point of autarky – potentially[13] one for each subcoalition. One important consequence of this difference is that the ccb-core need not be convex, and its center of gravity may not be one of it elements (see Example 3 in Section 5).

We can generalize this example to symmetric games with any number of players, as the following proposition shows. To do that, first we need to state an auxiliary condition that is sufficient – but not at all necessary, see an example later in this section – for the existence result.

Condition A (Increasing core center) $v_{\#}(k+1) - v_{\#}(k) \geq v_{\#}(k)/k$ *for* $k \in \{1, ..., n-2\}$.

Condition A is a strengthened version of both convexity and superadditivity for all subgames,[14] requiring that marginal contributions increase sufficiently fast in the size

[13] It could be the case that the cb-core of a coalition is empty, in which case it cannot credibly block any solution.

[14] It is straightforward to show that Condition A implies both superadditivity and convexity for any

of the coalition, similar to economies of scale. Technically, it ensures that the core center of a larger coalition give a higher payoff to the players. Note that we do not constrain the marginal increase to reach the grand coalition ($k = n-1$), which has important consequences as explained below.

Proposition 2 *Suppose (N, v) is symmetric, superadditive and satisfies Condition A. Then ccb-core(N, v) is non-empty, while the core(N, v) is non-empty only if we also have that $v_{\#}(n) - v_{\#}(n-1) \geq v_{\#}(n-1)/(n-1)$.*

Note the economic significance of this result. The core predicts that if the last member of the grand coalition does not add a large enough marginal gain to $v_{\#}(n-1)$, then his presence will prevent an agreement, because he can join blocking subcoalitions with some arbitrary agreements within them. When we restrict blocking to be consistent credible, expecting the core center as the outcome, an agreement within the grand coalition becomes feasible.[15]

Towards the proof of Proposition 2, we start with the characterization of the core and the core center of a symmetric game. While these results are straightforward, we provide their proofs for completeness.

Lemma 1 *If non-empty, the core of a symmetric game is invariant to relabeling of the players.*

Proof. Suppose the core of a symmetric game were asymmetric. Then there would exist an element of the core such that reshuffling the player indices, the assignment would no longer be in the core. But this is impossible as all the blocking coalitions are symmetric. ■

The following corollary is immediate.

subgame with $n-1$ members.

[15] On the other hand, when we have "convexity" for the entire range of coalition sizes, the core becomes non-empty in an interesting parallel with Moldovanu and Winter (1994), where convexity is a condition for non-cooperative implementation.

Corollary 1 *If it exists, the core center of a symmetric game distributes $v(N)$ equally among the players.*

Proof. Given that, by Lemma 1, the core is invariant to relabeling the players, the center of gravity must be as well. That is, each player has to earn the same utility. ■

We are now ready to prove the proposition.

Proof of Proposition 2. Our proof is constructive. We are going to demonstrate that, for some $\varepsilon \geq 0$, the assignment

$$\left(v_{\#}(1), v_{\#}(2)/2, ..., \frac{v_{\#}(n-1)}{n-1}, \frac{v_{\#}(n-1)}{n-1} + \varepsilon\right) \tag{1}$$

is in the ccb-core.[16] Start with observing that Condition A implies that the sequence of payoffs in (1) is weakly increasing. Now, by Corollary 1, the core center of each subgame of k players assigns $v_{\#}(k)/k$ to each member of the coalition. Consequently, since in (1) for no $k < n$ are there k payoffs strictly lower than $v_{\#}(k)/k$, the assignment cannot be credibly blocked. Thus, all we have left to prove is that the proposed assignment is feasible:

$$\frac{v_{\#}(n-1)}{n-1} + \sum_{1=1}^{n-1} \frac{v_{\#}(i)}{i} \leq v(N). \tag{2}$$

For $n = 2$ the inequality holds by superadditivity. Suppose $n > 2$ is even. Then we can split the terms in the sum of the left-hand side of (2) into three groups, $v_{\#}(1)+2\frac{v_{\#}(n-1)}{n-1}$, $\frac{2v_{\#}(n/2)}{n}$, and $\frac{v_{\#}(k)}{k} + \frac{v_{\#}(n-k)}{n-k}$ for $k \in \{2, ..., n/2-1\}$, where the last group is empty for $n = 4$. Next we show that

$$v_{\#}(1) + 2\frac{v_{\#}(n-1)}{n-1} \leq \frac{3v(N)}{n}. \tag{3}$$

Note that, by superadditivity,

$$\frac{3\left(v_{\#}(1) + v_{\#}(n-1)\right)}{n} \leq \frac{3v(N)}{n}.$$

Thus, a sufficient condition for (3) is

$$v_{\#}(1) + 2\frac{v_{\#}(n-1)}{n-1} \leq \frac{3\left(v_{\#}(1) + v_{\#}(n-1)\right)}{n}.$$

[16] Of course, any permutations of the player indices would lead to additional elements of the cb-core.

Rearranging and simplifying, we obtain

$$(n-1)v_{\#}(1) \leq v_{\#}(n-1),$$

what is again implied by superadditivity, and thus (3) is proven. Similarly,

$$\frac{2v_{\#}(n/2)}{n} \leq \frac{v(N)}{n}.$$

For $n = 4$, these two are the only groups, so the sum is no greater than $v(N)$, as required. For $n > 4$ we need to take into account the third group as well. Note that

$$\frac{v_{\#}(k)}{k} + \frac{v_{\#}(n-k)}{n-k} \leq \frac{2\left(v_{\#}(k) + v_{\#}(n-k)\right)}{n} \leq \frac{2v(N)}{n}, \tag{4}$$

where the second inequality again holds by superadditivity, while first inequality can be reduced to

$$\frac{v_{\#}(k)}{k} \leq \frac{v_{\#}(n-k)}{n-k}$$

what holds by Condition A. Summing the three upper bounds, noting that there are $n/2 - 2$ terms in the third group, we obtain

$$\frac{n/2+2}{n} v(N) < v(N)$$

what clearly holds for $n > 4$. Thus, we can add the difference between $v(N)$ and the sum to the last (highest) payoff (captured as ε in (1)), without affecting the blocking analysis above.

When $n > 2$ is odd, we do not have the middle term. We still have $v_{\#}(1) + 2\frac{v_{\#}(n-1)}{n-1}$, and the other terms are $\frac{v_{\#}(k)}{k} + \frac{v_{\#}(n-k)}{n-k}$ for $k \in \{2, ..., (n-1)/2\}$. We have $(n-3)/2$ of these pairs. Thus, when $n = 3$, we only have $\frac{3v(N)}{n} = v(N)$, as required. For greater odd n, using (3) and (4), we can derive the upper bound as

$$\frac{3v(N)}{n} + \frac{2v(N)}{n}(n-3)/2 = v(N).$$

Again, if there is a slack, the difference can be added to the highest payoff.

We only have left to prove that the core is non-empty only if additionally $v(N) - v_{\#}(n-1) \geq v_{\#}(n-1)/(n-1)$. Suppose otherwise. Consider blocking by coalitions of $n-1$ members. To avoid their success, any subset of N with $n-1$ members must earn at least $v_{\#}(n-1)$. The cheapest way to achieve this is assigning every player $v_{\#}(n-1)/(n-1)$. Consequently, we need $nv_{\#}(n-1)/(n-1) \leq v(N)$, which is exactly the condition stated in the proposition. ■

Before proceeding, let us display an example showing that Condition A – as well as superadditivity of the game – is far from necessary to have a non-empty ccb-core. Consider the four-player game with $v_{\#}(1) = 0$, $v_{\#}(2) = 2$, $v_{\#}(3) = 2$, $v_{\#}(4) = 3$. It clearly violates both Condition A and superadditivity. At the same time, $(0, 1, 1, 1)$ is in the ccb-core: the core center of any two-player coalition is $(1, 1)$, while the three-player coalitions have an empty core.

Our next result shows the non-emptiness of the ccb-core for three-player superadditive games, without any further restriction. This is in contrast to the well-known result that superadditivity is not sufficient for the core to be non-empty in three-player games, as the three-player subgames of our previous example illustrate.

Proposition 3 *Suppose (N, v) is superadditive. Then ccb-core(N, v) is non-empty for (two- or) three-player games.*

Proof. The two-player case is obvious. We will construct a payoff vector in ccb-core(N, v). It is immediate that $\mu\left(\{1,2\}, v_{12}\right) = \left(\frac{v(\{1,2\})+v(1)-v(2)}{2}, \frac{v(\{1,2\})-v(1)+v(2)}{2}\right)$.[17] We denote this vector by $(A12, B12)$, where the letters refer to the player whose payoff it is $(A = 1, B = 2, C = 3)$, and the numbers refer to the players in the coalition. Take highest deviation payoff of Player 1, $\max\{A12, A13\}$, and assign this payoff to her. This will ensure that no coalition including Player 1 will consistent-credibly block the imputation. For the rest of the assignations, choose $B23$ for Player 2 and $v(3)$ for player 3. Note that with these payoffs there is no credible blocking coalition: given that 1 is not willing to deviate, 2 could only deviate with 3, but he is paid off not to do that; finally 3 has no potential blocking partners, so we only need to give her

[17] Note that this coincides with the Nash Bargaining Solution.

her autarky payoff. Thus, all we have left to prove is that the budget of $v(1,2,3)$ is sufficient for these assignments. Suppose, without loss of generality, that $A12 \geq A13$. Then we have

$$A12 + B23 + v(3) = \frac{v(\{1,2\}) + v(1) - v(2)}{2} + \frac{v(\{2,3\}) + v(2) - v(3)}{2} + v(3)$$
$$= \frac{(v(\{1,2\}) + v(3)) + (v(1) + v(\{2,3\}))}{2} \leq v(\{1,2,3\}),$$

where the inequality follows by superadditivity. When the inequality is strict, we can arbitrarily distribute the surplus, $v(\{1,2,3\}) - A12 - B23 - v(3)$, among the players (adding it to the amount already assigned). ■

5 Characterization for three-player, superadditive games

Since the core and the ccb-core often differ for this class of games, while their calculations is still simple, the in depth analysis of three-player games is particularly instructive. 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. While not immediately obvious, our first observation is that it is not sufficient to consider blocking by the core center of two-person coalitions, as in some cases unilateral blocking may be more restrictive. Thus, our first step is to ensure Individual Rationality (IR): any solution in the ccb-core must be weakly greater than the autarky point (a, b, c). To simplify this task, let us write ccb-core$(\{1,2,3\}, v)$ as the sum of this point and the solution of the auxiliary game, $v' = (0, 0, 0; \alpha, \beta, \gamma; \alpha + \delta)$, where we have subtracted a, b, and c, from the payoffs of players 1, 2, and 3, respectively, in v.[18] Next, we turn to solving v', where we need not

[18] Since any credible blocking by a two-person coalition must be IR, no credible blocking can be less than (a, b, c). Thus, since we subtract the same amount from both sides of any inequality we consider

worry about IR beyond restricting attention to non-negative payoffs. We have that ccb-core($\{1,2\}, v'$) $= (w, \alpha - w, 0)$ for all $w \in [0, \alpha]$, ccb-core($\{1,3\}, v'$) $= (w, 0, \beta - w)$ for all $w \in [0, \beta]$, and ccb-core($\{2,3\}, v'$) $= (0, w, \gamma - w)$ for all $w \in [0, \gamma]$. Consequently, ccb-core-center($\{1,2\}, v'$) $= (\alpha/2, \alpha/2, 0)$, ccb-core-center($\{1,3\}, v'$) $= (\beta/2, 0, \beta/2)$ and ccb-core-center($\{2,3\}, v'$) $= (0, \gamma/2, \gamma/2)$. Thus, ccb-core($\{1,2,3\}, v'$) is given by the subset of the simplex $x + y + z = \alpha + \delta$,[19] that is not strictly below any of the three ccb-core-centers just identified. It is easiest to get a feel for the set graphically. We are going to project the simplex on to the 1-2 plane (looking from above, when the vertical dimension is Player 3's payoff).

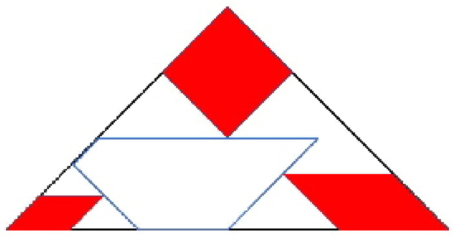

Figure 1

The white "cubist clover" in the middle represents the agreements in the ccb-core. The square on top contains the agreements removed by the credible block of the coalition formed by Players 1 and 2. These are the ones that would give Player 3 "too much" (the upper vertex is $(0, 0, \alpha + \delta)$).[20] The two parallelograms are the equivalent sets for the other two coalitions.

The blue-bordered pentagonal area in the middle represents the core. Its edges that are not on the edges of the simplex are parallel to the farthest edge of the simplex and pass through the interior vertex of the corresponding parallelogram.

to evaluate whether a payoff vector is in the cb-core, the claimed equivalence holds.

[19] That is, the triangle defined by $[(\alpha + \delta, 0, 0), (0, \alpha + \delta, 0), (0, 0, \alpha + \delta)]$.

[20] The other three vertices of the square are $(\alpha/2, 0, \alpha/2 + d)$, $(\alpha/2, \alpha/2, d)$ and $(0, \alpha/2, \alpha/2 + d)$.

Note that the parallelograms may overlap[21] but only in the interior of the simplex, on the edges they can at most "touch". In fact, none of them can grow beyond half the length of an edge (by superadditivity). As the edges (including their vertices) of the parallelograms in the interior of the simplex are still immune to credible blocking, even when the whole simplex is covered, the ccb-core is not empty, see Example 2 below.

Let us illustrate the solution with some examples, which showcase the main economic intuitions we can gain.

Example 1 For $v = (0,0,0;1,1,0;1+\delta)$, 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. The core-center is $(1,\delta/2,\delta/2)$.

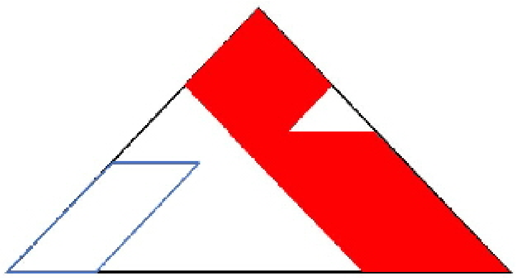

Figure 2

To calculate ccb-core$(\{1,2,3\},v)$, note that

$$\begin{aligned} \text{core}\left(\{1,2\},v_{(1,2)}\right) &= \text{core}\left(\{1,3\},v_{(1,3)}\right) = (x,1-x) \text{ for all } x \in [0,1], \\ \text{while core}\left(\{2,3\},v_{(2,3)}\right) &= (0,0). \end{aligned}$$

Consequently,

$$\begin{aligned} \text{core-center}\left(\{1,2\},v_{(1,2)}\right) &= \text{core-center}\left(\{1,3\},v_{(1,3)}\right) = (1/2,1/2) \text{ and} \\ \text{core-center}\left(\{2,3\},v_{(2,3)}\right) &= (0,0). \end{aligned}$$

[21] Note that the rhombi must overlap when the core is empty (but the reverse is not true).

This means that the ccb-core consists of two – for $\delta > 1/2$ overlapping – components (triangles): in one we avoid blocking by giving Player 1 at least 1/2, in the other we give both Players 2 and 3 at least 1/2.

$$\begin{aligned} \text{ccb-core}(\{1,2,3\},v) &= (x,y,z) \text{ with } x \geq 1/2,\ x+y+z = 1+\delta, \text{ and } y,z \geq 0 \\ \cup (x,y,z) \text{ with } y &\geq 1/2,\ z \geq 1/2,\ x+y+z = 1+\delta, \text{ and } x \geq 0. \end{aligned}$$

Finally, the ccb-core-center$(\{1,2,3\},v)$ is a point on the segment connecting the centers of each component $(2/3+\delta/3, 1/6+\delta/3, 1/6+\delta/3)$ – what is the Shapley value of $(\{1,2,3\},v)$ – and $(\delta/3, 1/2+\delta/3, 1/2+\delta/3)$. When $\delta = 0$, the second component collapses onto the point $(0,1/2,1/2)$, and thus has no weight in the averaging, so the ccb-core-center becomes $(2/3,1/6,1/6)$, the Shapley value. As δ increases, it moves closer to the other center, but it always stays closer to the first one.

The core is usually defended here as if Player 1 could organize an auction. But that is not the idea behind the core. Instead, it has the following implementation: An potential deal (x,y,z), is put on the table in front of the players. Any player can decide to make a take-it-or-leave-it offer to another one to share their value. Say $y > 0$, then Player 1 can offer Player 3 $z + \frac{y}{2}$ and keep $x + \frac{y}{2}$, blocking the putative solution. Similarly for $z > 0$. Note the crucial detail: when z is small, Player 1 has to offer Player 3 little, while if z is large she has to offer a lot: the negotiation between the two needs to have a different outcome depending on the offer on the table. This means that they must negotiate before the three-way deal disappears.

In case of ccb-core, if Player 1 wants to block with Player 2, all she is allowed to do is to threaten to take Player 2 away and negotiate with him, giving up on the three-way negotiation. In expectation, this two-way negotiation leads to equal sharing. Thus, Player 1 cannot get more than 1/2 following the blocking attempt. As a result, she cannot block allocations that give her half or more. At the same time, the most enticing offer she can make to another player is also 1/2. Thus, if both 2 and 3 already have 1/2 or more in the putative deal, she cannot block it with either of them. This, second component captures the possibility collusion between Players 2 and 3. By refusing to compete with each other they can extract surplus from Player 1.

Example 2 Let $v' = (0,0,0;1,1,1;1)$. We set $\delta = 0$ for simplicity. Now the core is empty, and thus there is no core-center. On the other hand,

$$\text{core-center}\left(\{1,2\}, v'_{(1,2)}\right) = \text{core-center}\left(\{1,3\}, v'_{(1,3)}\right) = \text{core-center}\left(\{2,3\}, v'_{(2,3)}\right) = (1/2, 1/2).$$

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:

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

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. Finally, removing the uncertainty we obtain the ccb-core-center$(\{1,2,3\}, v') = (1/3, 1/3, 1/3)$. Same as the Shapley value.

Example 3 Let $v'' = (0,0,0;1,0,0;1)$. Now the core is $(x, 1-x, 0)$, $x \in [0,1]$, and the core-center is $(1/2, 1/2, 0)$. Same as the Shapley value. On the other hand,

$$\begin{aligned}
\text{core-center}\left(\{1,2\}, v''_{(1,2)}\right) &= (1/2, 1/2) \\
\text{core-center}\left(\{1,3\}, v''_{(1,3)}\right) &= \text{core-center}\left(\{2,3\}, v''_{(2,3)}\right) = (0,0),
\end{aligned}$$

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

the two white triangles. Finally, ccb-core-center$(\{1,2,3\}, v'') = (5/12, 5/12, 1/6)$.

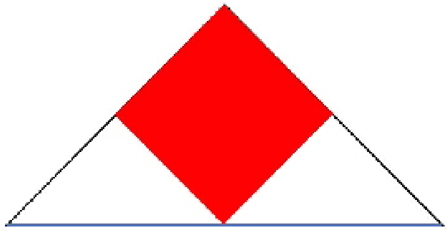

Figure 3

Note that 3 is a dummy player – his presence does not change the value of any coalition he joins – but he has a positive payoff. Is this a problem, or an insight? We believe that while from a normative point of view one might argue that a dummy player should earn their autarky value, from a positive standpoint it may be wrong to ignore their presence. As long as either 1 or 2 earns more than 1/2, they cannot block the agreement. It is as if 3 bribed one of them in order to get a positive payoff.

It is interesting to note that the ccb-core-center is not in the ccb-core. Indeed, core-center$\left(\{1,2\}, v''_{(1,2)}\right)$ strictly dominates the ccb-core-center for Players 1 and 2.

6 Concluding remarks

We have presented a modified version of the classical solution concept of the core, based on consistently requiring blocking threats to represent the expected outcome of the blocking coalition "going it alone". The new solution is a core extension, which is non-empty for empirically relevant situations where the core is empty.

The apparent softening of competition that we have seen in our examples, as a result imposing credibility on blocking, gives rise to interesting – and empirically relevant – phenomena that are absent in the core. In particular, players that have little to contribute may appropriate a disproportionately large proportion of the surplus, when concurrently other players are also earning a lot (enough to discourage them from blocking the deal). As a result, a coalition of players that can create little value on their own, may appropriate most of the surplus by "colluding". While this is perhaps unfair, collusion is a real life phenomenon, so our solution would better be able to account for it. Consider the case of an entrepreneur and two workers. The former only needs one of the workers to create value. We might think that the firm can extract most of the surplus buy making the workers compete. But what if the workers form a union?

A possible explanation for this phenomenon is the misalignment between how a coalition's value is allocated internally and the marginal contributions of its members to alternative coalitions. Intuitively, one might expect that collusion would be facilitated if the prospective blocking coalition offered a more favorable allocation to the pivotal player. However, the opposite holds. Collusion becomes more difficult when the pivotal player receives a relatively small share of the blocking coalition's surplus. In such cases, the remaining member(s) capture a larger portion of the surplus and therefore have stronger incentives to deviate, making the blocking coalition more likely to form.

An interesting open question is how – and whether – to extend the consistent credibility requirement recursively to all subcoalitions. This would require combining set-valued and point-valued solutions in sequence. A logical way to proceed could be the following: We begin with two-player coalitions, wich only face unilateral blocking, what is always credible. We compute the expected outcomes within these cores, that represent what players can expect if they deviate in pairs. For three-player coalitions, the relevant credible threats are the above calculated core centers, together with autarky. We identify the set of allocations that are immune to these threats and again take their centers of gravity. Iterating this procedure, we successively construct expectations for larger coalitions, until we obtain a set-valued solution for the grand coalition.

This method can be formalized in a straightforward manner, and the corresponding solution can be computed. However, the approach suffers from an important shortcoming: it fails to respect credibility, at least as defined above. Specifically, because the ccb-core center need not belong to the ccb-core itself (see Example 3 in Section 5), using the ccb-core center as a blocking threat does not guarantee that we select an element of the subgame's solution (i.e., the ccb-core). Even if one sets aside the conceptual difficulty arising from the lack of convexity – perhaps arguing that the ccb-core center still represents an "expected outcome" despite not being feasible – the resulting behavior of the solution remains somewhat irregular.

References

[1] 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.

[2] Burguet, Roberto, and Ramon Caminal. "Coalitional bargaining with consistent counterfactuals." *Journal of Economic Theory* 187 (2020): 105013.

[3] Dutta, Bhaskar, Debraj Ray, Kunal Sengupta, and Rajiv Vohra. "A consistent bargaining set." *Journal of Economic Theory* 49(1) (1989): 93-112.

[4] Edgeworth, Francis Ysidro. *Mathematical psychics: An essay on the application of mathematics to the moral sciences.* Kegan Paul, London (1881).

[5] Gillies, Donald B. "Some theorems on n-person games." Doctoral Thesis (1953). Department of Mathematics, Princeton.

[6] Gillies, Donald B. "Solutions to general non-zero-sum games." *Contributions to the Theory of Games* 4, no. 40 (1959): 47-85.

[7] 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.

[8] Harsanyi, John C., and Reinhard Selten. *A general theory of equilibrium selection in games*. MIT Press Books, Cambridge MA (1988).

[9] Kalai, Ehud, Andrew Postlewaite, and John Roberts. "A Group Incentive Compatible Mechanism Yielding Core Allocation." *Journal of Economic Theory* 20 (1979): 13-22.

[10] Moldovanu, Benny, and Eyal Winter. "Core implementation and increasing returns to scale for cooperation." *Journal of Mathematical Economics* 23(6) (1994): 533-548.

[11] Okada, Akira, and Eyal Winter. "A Non-cooperative Axiomatization of the Core." *Theory and Decision* 53(1) (2002): 1-28.

[12] 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.

[13] Platt, Brennan C. "Spoilers, blocking coalitions, and the core." *Social Choice and Welfare* 33(3) (2009): 361-381.

[14] Ray, Debraj. "Credible coalitions and the core." *International Journal of Game Theory* 18 (1989): 185-187.