ISER Seminar Series
4-6 pm, Thu, September 21, 2017

COOPERATIVE STRATEGIC GAMES

ELON KOHLBERG AND ABRAHAM NEYMAN

Abstract. The value is a solution concept – developed by Shapley, Nash, and Harsanyi – that reflects both the cooperative and competitive aspects of strategic games. The value provides an a priori evaluation of the position of each player, in particular one whose economic worth derives from an ability to inflict losses on other players, e.g., a municipal employee who can deny or delay building permits. Applications of the value in economics have been rare, at least in part because the existing definition requires a complex scheme that is not justified by any set of basic assumptions and that does not easily lend itself to computation in specific models. We provide an axiomatic characterization, i.e., we prove that the value is the unique function, from n-player strategic games to n-dimensional vectors of payoffs, that satisfies a short list of desirable properties. We also provide a formula for computing the value.

1. Introduction

A *strategic game* is a general model for a competitive interaction. Each player selects a strategy, and the combined choices of all the players determine a payoff to each of them.

The *"Shapley program,"* implicitly outlined in [13], is to define a solution concept that describes the a priori *value* of each position in a strategic game. The importance of the program cannot be overstated. As Shapley [13] wrote: "The possibility of evaluating the prospect [of playing a game] is of crucial importance to the successful application of [game] theory."

The value ought to reflect both the cooperative and the competitive aspects of a strategic game. One may think of it as the expected payoff in a cooperative process that takes into account all the players' strategic possibilities, including their capacity to make *threats* against one another.

We make the simplifying assumption that *utility is transferable*, i.e., that the players' payoffs are measured in units of a commodity that is freely exchangeable, like money. Therefore it is reasonable to expect that players will coordinate their strategic choices to maximize the sum of their payoffs, and that *side payments* will be made in accordance with the different threat powers of the different players.

A value solution provides an a priori assessment of the side payments. Thus it is a powerful tool for studying a variety of economic phenomena where side payments are made in response to explicit or implicit threats, e.g., corruption and bribery. Of course, the

Date: August 20, 2017.

extent to which this tool can be used for the analysis of real-world phenomena depends on the extent to which the definition of the value captures the impact of threats.

Shapley [14] provided the original definition, and Harsanyi [5] suggested a modification. We concur with Harsanyi's view that his definition is preferable. Its essential advantage is that it takes into account the potential damage of a threat not only to the threatened party but also to the party making the threat. This is crucially important for understanding the behavior of "spoilers," players whose economic value derives from their ability to inflict losses on other players, e.g., a municipal employee who is in a position to arbitrarily deny building permits. The value solution ought to exhibit a reduction in the economic value of a spoiler when the act of spoiling is costly. In the case of the municipal employee, the relevant costs might be probabilistic expectations of a fine or a prison term.

Harsanyi calls his solution the modified Shapley value; others call it the Harsanyi–Shapley value; we call it simply *the value*.

Despite being a prime tool for the analysis of important phenomena, the value solution for strategic games has not had a major impact on economics. Outside the work of Aumann and Kurtz [1, 2] and Aumann, Kurtz, and Neyman [3, 4], there have been few, if any, applications. One reason, perhaps, is that the available definition, which we describe in Section 11, is rather complex. We remedy this by providing a transparent definition. But a more important reason is that the available definition, which amounts to an algorithmic procedure, does not make clear what assumptions must be made in order to justify the solution.

The main result of this paper is an axiomatic characterization of the value of strategic games. That is, we prove that the value is the unique function, from n-person strategic games to n-dimensional vectors of payoffs, that satisfies a short list of axioms. The axioms elucidate the underlying assumptions and make it possible to judge the reasonableness of each application in its specific context.

There are five axioms. Four of these are relatively standard, being the strategic-game analogs of the classic Shapley axioms for the value of coalitional games. The fifth axiom rests on the notion of "balanced threats." It is motivated by the observation that, since the sum of the payoffs is fixed, any demand for a payoff by a player or a group of players must come, of necessity, at the expense of the other players. Thus, we consider the two-player zero-sum game between a subset S of the set of players, N, and its complement, $N \setminus S$, where the players in each of these subsets coordinate their strategies to act as a single player, and where the payoff to player S is the difference between the sum of the (original game) payoffs to the players in S and the sum of the payoffs to the players in $N \setminus S$; and we define $d(S)$ as the minmax value of this game. We think of $d(S)$ as representing the *threat power* of the players in S. The axiom of balanced threats says that if the threat powers of all subsets are the same then all players receive the same payoff. An alternative formulation of the axiom is that, if the threat power of every subset is zero, then all players receive the same payoff. The equivalence of the two formulations follows from the minmax theorem, which says that $d(S) = -d(N \setminus S)$ and therefore that $d(S) = 0$ when $d(S) = d(N \setminus S)$.

An auxiliary result is a formula for the computation of the value. The formula says that the value of a player in an n-person strategic game is an average of the threat powers, $d(S)$, of the subsets of which the player is a member. Specifically, if $\delta_{i,k}$ denotes the average of $d(S)$ over all k-player subsets that include i, then the value of player i is the average of $\delta_{i,k}$ over $k = 1, 2, \ldots, n$.

We wish to emphasize that the fact that the value of a strategic game depends only on the threat powers $d(S)$ is not an assumption but rather a conclusion. Indeed, the key step in proving the main result is the derivation of this conclusion from the axioms. And, as the axiom of balanced threats plays a key role, we wish to emphasize that it imposes restrictions on the value function only in rare games where the threat powers of all subsets are equal.

We apply the formula in some simple examples. In one example, the economic output of a large number of individuals is predicated on the approval of a regulator. Computation of the value indicates that the regulator's position is worth 25% of the total output; and if approval is required from two regulators, then their combined positions are worth a full 42% of the total output. However, if approval is required from only one of the two regulators, then their combined positions are worth just 8.5% of the output. In another example, it is assumed that making good on an implicit threat to deny approval (for no valid reason) exposes the regulator to potential punishment. If the expected cost of the punishment is a fraction c of the lost output, then the value of the regulator's position decreases by a factor of $(1 - c)^2$. Such examples quantify the strength of the temptation to use the power of approval in order to extract side payments.

The next section is best skipped in a first reading. It discusses the historical development of the ideas, which is rather involved. The rest of the paper is organized as follows. In Sections 3 and 4 we define the axioms and state the main results: the axiomatic characterization and the formula for computing the value. In Section 5 we apply the formula in a number of examples. Sections 6 and 7 provide the background on games of threats and present an alternative definition of the value in terms of such games. Section 8 presents preliminary results, some of which are of interest in their own right, and Section 9 presents the proof of the main theorem. In Section 10 we provide a characterization of the von Neumann–Morgenstern and Shapley value that parallels the characterization of the value. In Section 11 we describe the n-player generalizations of Nash's "bargaining with variable threats" by Harsanyi and by Myerson, and in Section 12 we discuss the coco value of Kalai and Kalai. In the Appendix we provide additional properties of the value and show that all the axioms are tight; i.e., if any of them is dropped then the uniqueness theorem is no longer valid.

2. History of the Concepts

In the classic work [16] of von Neumann and Morgenstern (VNM), the starting point for the cooperative analysis of strategic games is to reduce every such game to a characteristic function, nowadays called a *coalitional game*, which assigns to every subset of players ("coalition") S a single number, $v(S)$, defined as the total payoff that the members of S

can guarantee, i.e, the maxmin of the sum of the payoffs to the members of S, where the max is over all the correlated strategies of S and the min is over the correlated strategies of the complement of S. Having reduced strategic games to coaltional games, VNM focused on developing their solution concept for coalitional games, the "stable set."

In contrast to VNM's set-valued solution, Shapley highlighted the need to define a single-valued function that assigns to each strategic game a vector of payoffs, representing the value of each role in the game. Shapley accepted the VNM approach of reducing strategic games to coalitional games; thus he addressed the problem of defining a value function for coalitional games. In a seminal paper [14] he formulated properties ("axioms") that would be desirable in such a function and proved that – remarkably – a mere four of them uniquely imply one particular function, the "Shapley value."

It would seem, then, that Shapley's goal of defining a value function for strategic games had been accomplished: given a strategic game, transform it to its VNM coalitional form, then apply the Shapley value. But there were doubts. The doubts, centering on the adequacy of the VNM coalitional game, were raised by VNM[1] and Shapley[2] themselves, as well as by Luce and Raiffa [8], Harsanyi [5], and Myerson [10].

We shall discuss these concerns shortly but first we wish to describe a development that pointed the way to a resolution of the difficulties. Nash [11] pioneered the notion of a cooperative solution for strategic games. He defined such a concept for two-person games and proved an existence and uniqueness theorem. The solution is derived by means of "bargaining with variable threats." In an initial competitive stage, each player declares a "threat" strategy, to be used if negotiations break down; the outcome resulting from deployment of these strategies constitutes a "disagreement point." In a subsequent cooperative stage, the players coordinate their strategies to achieve a Pareto optimal outcome, and share the gains relative to the disagreement point; the sharing is done in accordance with principles of fairness. Nash proves that in this game, where each player chooses a threat strategy and the resulting payoffs are as described above, there is a unique Nash equilibrium payoff.

As was pointed out by Shapley [15], when players can transfer payoffs among themselves the outcome of "bargaining with variable threats" can be described very simply, as follows.[3] Let s denote the maximal sum of the players' payoffs in any entry of the payoff matrix, and let d be the minmax value of the zero-sum game constructed by taking the difference

[1] von Neumann and Morgenstern [16] wrote: "In a general [-sum] game the advantage of one group of players need not be synonymous with the disadvantage of the others. In such a game moves – or rather changes in strategy – may exist which are advantageous to both groups. . . . Does our approach not disregard this aspect?"

[2] Shapley [13] wrote: "Serious doubt has been raised as to the adequacy with which the characteristic function describes the strategic possibilities of a general-sum game. The difficulty, intuitively, is that the characteristic function does not distinguish between threats that damage just the threatened party and threats that damage both parties. This criticism, however, does not apply with any force to the constant-sum case.

[3] Kalai and Kalai [6] independently discovered Shapley's reformulation of "bargaining with variable threats" in two-person games and named the resulting solution concept the *coco value*. Their main result is an axiomatic characterization of the coco value. We discuss it in Section 12.

between player 1's and 2's payoffs. Then the Nash variable-threats solution splits the amount s in such a way that the difference in payoffs is d. Specifically, the payoffs to players 1 and 2 are, respectively,

$$\frac{1}{2}s + \frac{1}{2}d \text{ and } \frac{1}{2}s - \frac{1}{2}d. \tag{1}$$

We now return to the difficulties with the VNM coalitional game. Consider the two-person game below (Luce and Raiffa [8], Section 6.8).

Example 1.

$$\begin{bmatrix} 2,\ 1 & -1,-2 \\ -2,-1 & 1,\ \ 2 \end{bmatrix}.$$

At first blush the game looks entirely symmetrical. The set of feasible payoffs (the convex hull of the four entries in the matrix) is symmetrical; and the maximum payoff that each player can guarantee is the same, namely, 0. Thus the VNM coalitional game is symmetrical: $v(1) = v(2) = 0$, $v(1, 2) = 3$, and its Shapley value is $(1.5, 1.5)$.

However, the Nash analysis reveals a fundamental asymmetry. In the zero-sum game of differences

$$\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$$

the minmax value is not zero but rather $d = 1$. This indicates that the threat power of player 1 is greater than the threat power of player 2. The Nash variable-threats solution reflects this advantage. It is $(1.5 + .5,\ 1.5 - .5) = (2, 1)$.

The difficulty with the VNM coalitional game – the reason that it does not properly capture the threat powers – arises because a coalition is allowed to deploy two different strategies, one for maximizing its own payoff and the other for minimizing the complementary coalition's payoff. In the example, player 2 plays $(\frac{2}{3}, \frac{1}{3})$ in order to maximize his own payoff, but plays $(\frac{1}{3}, \frac{2}{3})$ in order to minimize player 1's payoff.

Harsanyi [5] proposed a modification in the definition of the value. It is this modification that we now call "the value." We describe Harsanyi's method in Section 11. Here we provide a simpler description, as follows.

Instead of considering two separate zero-sum games between a coalition S and its complement, one that focuses on the payoff to S and the other on the payoff to $N \setminus S$, we consider a single game that focuses on the *difference* between these payoffs; and we assign to each coalition S a single number, $d(S)$, defined as the maximal difference between the total payoffs to S and to $N \setminus S$ that the members of S can guarantee.

Now d is not a coalitional game. It may fail to satisfy the single condition required of a set function to qualify as a coalitional game, namely $d(\emptyset) = 0$. This condition is essential for the formula of the Shapley value, which assigns to each player i an average of his marginal contributions, including the marginal contribution $v(i) - v(\emptyset) = v(i)$. However, we show ([7]) that an appropriate modification of the definition of the Shapley value applies to set functions such as d, which satisfy the condition that $d(S) = -d(N \backslash S)$ for all S, and which we call "games of threats." The value of the strategic game is then obtained by taking the Shapley value of d. We refer to this modification by Harsanyi of the VNM–Shapley value as *the value* of a strategic game.

It is easy to verify that the value coincides with the Nash variable-threats solution in two-player games and that the value coincides with the VNM–Shapley value in constant-sum games and in games without externalities; see Remark 10 and Proposition 12.

We wish to emphasize that neither Shapley's original definition of a value for strategic games nor Harsanyi's modification rest on an axiomatic foundation, as the first step – that of reducing the strategic game to a coalitional form – is arbitrary.

We end this section with an example that demonstrates the response of the two notions of value to an increase in the cost of making a threat.

Example 2.

$$\begin{bmatrix} 2,4 & 3,3 \\ 0,0 & 0,0 \end{bmatrix}.$$

Purely-competitive analysis considers player 1's "down" strategy to be an *incredible threat*; thus the availability of this strategy does not affect the equilibrium outcome, $(2,4)$. But in an explicit or implicit cooperative process, player 1's option to play "down" is a source of power. Indeed, both the VNM–Shapley value and the value exhibit side payments from player 2 to player 1.

The VNM coalitional game is $v(1) = 2$, $v(2) = 0$ and $v(1,2) = 6$. Thus the VNM–Shapley value is $(4,2)$.

The minmax value of the game of differences,

$$\begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix},$$

is 0. Thus the value, i.e., the Nash variable-threats solution, is $(3,3)$.

Now consider the following variant.

Example 3.

$$\begin{bmatrix} 2,4 & 3,3 \\ -1,0 & 0,0 \end{bmatrix}.$$

The security levels of the players are unchanged. Thus the VNM coalitional game is unchanged and the VNM–Shapley value is still $(4,2)$. But the minmax value of the game of differences,

$$\begin{bmatrix} -2 & 0 \\ -1 & 0 \end{bmatrix},$$

is now -1; thus the value is $(2.5, 3.5)$: The increased cost of the threat has had an impact.

3. The Axioms

A strategic game is a triple $G = (N, A, g)$, where

- $N = \{1, \ldots, n\}$ is a finite set of players,
- A^i is the finite set of player i's pure strategies, and $A = \prod_{1=1}^{n} A^i$,
- $g = (g^i)_{i \in N}$, where $g^i \colon A \to \mathbb{R}$ is player i's payoff function.[4]

We use the same notation, g, to denote the linear extension

- $g^i \colon \Delta(A) \to \mathbb{R}$,

where for any set K, $\Delta(K)$ denotes the probability distributions on K, and we denote

- $A^S = \prod_{i \in S} A^i$, and
- $X^S = \Delta(A^S)$ (correlated strategies of the players in S).

We define the *direct sum* of strategic games as follows.[5]

Definition 1. *Let $G_1 = (N, A_1, g_1)$ and $G_2 = (N, A_2, g_2)$ be two strategic games. Then $G := G_1 \oplus G_2$ is the game $G = (N, A, g)$, where $A = A_1 \times A_2$ and $g(a) = g_1(a_1) + g_2(a_2)$.*

Remark 1. The game $G_1 \oplus G_2$ models a situation where the same set of players simultaneously play two games that are independent, i.e., where the moves in one game do not influence the other game.

Remark 2. It is easy to verify that the operation $\oplus$ is, informally, commutative and associative.[6] However, there is no natural notion of inverse. (In general, $G \oplus (-G) \neq 0$.)

Denote by $\mathbb{G}(N)$ the set of all n-player strategic games. Let $\gamma : \mathbb{G}(N) \to \mathbb{R}^n$. This may be viewed as a map that associates with any strategic game an allocation of payoffs to the players. We consider a list of axioms on γ. To that end we first introduce a few definitions.

Let $G \in \mathbb{G}(N)$. We define the *threat power* of coalition S as follows:[7]

$$(\delta G)(S) := \max_{x \in X^S} \min_{y \in X^{N \setminus S}} \left(\sum_{i \in S} g^i(x, y) - \sum_{i \notin S} g^i(x, y) \right), \tag{2}$$

and we say that S has an *effective threat* if $(\delta G)(S) > 0$.

We say that *i and j are substitutes in G* if $A^i = A^j$ and $g^i = g^j$; and for any $a, b \in A^N$, if $a^i = b^j, a^j = b^i$, and $a^k = b^k$ for all $k \neq i, j$, then $g(a) = g(b)$.

[4]The assumption that the sets of players and strategies are finite is made for convenience. The results remain valid when the sets are infinite, provided the minmax value exists in the two-person zero-sum games defined in the sequel.

[5]von Neumann and Morgenstern [16], Section 27.6.2, refer to this operation as the *superposition* of games.

[6]Formally, $G_1 \oplus G_2$ is not the same game as $G_2 \oplus G_1$, because $A_1 \times A_2 \neq A_2 \times A_1$.

[7]Expressions of the form max or min over the empty set should always be ignored.

We say that i is a *null player* in G if $g^i(a) = 0$ for all a; and if $a^k = b^k$ for all $k \neq i$, then $g(a) = g(b)$.

Our list of axioms is as follows. For all strategic games G,

- **Efficiency** $\sum_{i \in N} \gamma_i G = \max_{a \in A^N} (\sum_{i \in N} g^i(a))$.
- **Balanced threats** If no subset of players has an effective threat (equivalently, if $\delta G(S) = 0$ for all S) then $\gamma_i G = \gamma_j G$ for all $i, j \in N$.
- **Symmetry** If i and j are substitutes in G then $\gamma_i G = \gamma_j G$.
- **Null player** If i is a null player in G then $\gamma_i G = 0$.
- **Additivity** $\gamma(G_1 \oplus G_2) = \gamma G_1 + \gamma G_2$.
- **Individual rationality** $\gamma_i(G) \geq \max_{x \in X^i} \min_{y \in X^{N \setminus i}} g^i(x, y)$.

Efficiency says that the players are allocated the maximum available payoff.[8]

The axiom of balanced threats says that if no player – regardless of the additional players who have joined him – can effectively threaten the remaining players, then all players receive the same amount. (The equivalent formulation follows from the minmax theorem, which says that $(\delta G)(S) = -(\delta G)(N \setminus S)$.)

Symmetry says that players whose payoffs are identical everywhere, and whose strategies can be switched without impacting any payoff, receive the same allocation.

The null-player axiom says that a player whose actions do not affect any player's payoff, and whose own payoff is identically zero, receives an allocation of zero.

Additivity says that if the payoff to the players is the sum of their payoffs in two games that are unrelated to each other then the allocation to the players is the sum of their allocations in these two games. Note that the purely competitive notion of Nash equilibrium payoffs, viewed as a set function, also satisfies additivity.

Individual rationality says that each player receives at least his security level, i.e., the maximal payoff that the player can guarantee unilaterally, irrespective of the strategies of the other players.

Our main result is that *there exists a unique map from $\mathbb{G}(N)$ to $\mathbb{R}^n$ satisfying the axioms of efficiency, balanced threats, symmetry, null player, and additivity*, and that this map satisfies individual rationality.

Remark 3. The expression "threat power" does not make much sense when applied to the set N of all players. The condition that $(\delta G)(N) = 0$ simply means that the maximal sum of payoffs is zero. Thus the axiom of balanced threats can be stated as follows. If all *proper* subsets S (i.e., $S \neq \emptyset$ and $S \neq N$) have the same threat power *and* the total available payoff is zero then all players receive the same payoff. A less restrictive version of the axiom (that would imply a weaker uniqueness theorem) drops the second restriction: if $(\delta G)(S) = 0$ for all proper subsets then $\gamma_i G = \gamma_j G$ for all $i, j \in N$.

[8]Efficiency seems to be a reasonable axiom for the evaluation of a cooperative outcome. But one can imagine models where this axiom is rejected. Important examples are Ray and Vohra [12] and Maskin [9].

Remark 4. Another less restrictive version of the axiom of balanced threats is as follows. If the function d is symmetrical, i.e., $d(S)$ is the same for all coalitions S with the same number of players, then all players receive the same payoff.

Remark 5. There are many additional desirable properties of the value that we do not assume but rather deduce from the axioms. These include dependence on the reduced form of the game (removing strategies that are convex combinations of other strategies does not affect the value), homogeneity of degree one ($\gamma(\alpha G) = \alpha\gamma G$ for $\alpha > 0$), time-consistency ($\gamma(\frac{1}{2}G_1 \oplus \frac{1}{2}G_2) = \frac{1}{2}\gamma G_1 + \frac{1}{2}\gamma G_2$; i.e., it does not matter if an allocation is determined before or after the resolution of uncertainty about the game), monotonicity in actions (removing a pure strategy of a player does not increase the player's value),[9] independence of the set of players (addition of null players does not affect the value of the existing players), shift-invariance (adding a constant payoff to a player increases the player's value by that constant), a stronger form of symmetry (the names of the players do not matter), and continuity ($\gamma(G_n) \to \gamma G$ whenever $G_n = (N, A, g_n), G = (N, A, g)$, and $g_n \to g$).

Remark 6. We do not require, nor are we able to deduce, that $\gamma(\alpha G) = \alpha\gamma G$ for negative α. Such a requirement, which is natural in the context of coalitional games, would make no sense in the context of strategic games. The game $-G$ involves dramatically different strategic considerations than the game G, and so there is no reason to expect a simple relationship between the values of the two games.

4. The main result

Our main result is as follows.

Theorem 1. *There is a unique map from $\mathbb{G}(N)$ to $\mathbb{R}^n$ that satisfies the axioms of symmetry, null player, efficiency, balanced threats, and additivity. It may be described as follows.*

$$\gamma_i G = \frac{1}{n}\sum_{k=1}^{n} \delta_{i,k}, \tag{3}$$

where $\delta_{i,k}$ denotes the average of the $(\delta G)(S)$ over all k-player coalitions that include i. Furthermore, this map satisfies the axiom of individual rationality.

We shall refer to the above map as the *value* for strategic games.

Remark 7. The uniqueness theorem and formula (3) apply to Bayesian games, as these are special cases of strategic games.

Remark 8. There is only one n-player coalition, namely, N. Thus $\delta_{i,n} = \max_{x\in X^N}(\sum_{j\in N} g^j(x))$ is the maximum total payoff.

[9] These four properties follow from formula (3) and the corresponding properties of the minmax value of zero-sum games.

Remark 9. Formula (3) implies that the value of G is a function of the threats, $(\delta G)(S)$.

Remark 10. In a two-player game, the value coincides with the Nash variable-threats solution. Formula (3) becomes $\gamma_1 G = \frac{1}{2}\delta G(1) + \frac{1}{2}\delta G(1,2)$, which is the same as (1).

Remark 11. Each player is allocated a weighted average of the $\delta G(S)$ over the coalitions S that include that player. The weight is the same for all coalitions of the same size but different for coalitions of different size. Specifically, for each $k = 1, \ldots, n$, the total weight of $\frac{1}{n}$ is divided among the $\binom{n-1}{k-1}$ coalitions of size k that include i. Thus the formula can be rewritten as follows.

$$\gamma_i G = \frac{1}{n} \sum_{k=1}^{n} \frac{1}{\binom{n-1}{k-1}} \sum_{\substack{S: i \in S \\ |S| = k}} \delta G(S). \tag{4}$$

The proof of Theorem 1 requires the notion of games of threats [7]. We provide the relevant definitions and results in Section 6.

5. Examples

In each of the examples below we apply formula (3) to determine the value. These are toy examples, but they illustrate that the value can be a measure of the economic worth of a public official who has the authority to make decisions in matters of financial importance to private individuals or companies, and who weighs the benefits of receiving bribes against the costs of potential penalties. Thus the notion of value provides a tool for studying the effectiveness of alternative hierarchies of decision-making authority and the associated supervisory and disciplinary regimes.

Example 4. This is a three-player game. Player 1 chooses the row, player 2 chooses the column, and player 3 has only a single strategy. The payoff matrix is[10]

$$\begin{bmatrix} 2,2,2 & 0,0,0 \\ 0,0,0 & 1,1,1 \end{bmatrix}.$$

Now, letting $d = \delta G$,

$$d(1) = \text{minmax} \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix} = -\frac{2}{3},$$

$$d(1,3) = \text{minmax} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = \frac{2}{3},$$

and $d(1,2) = \max(2,0,1) = 2$, and $d(1,2,3) = \max(6,0,3) = 6$.

Thus $\gamma_1 = \frac{1}{3} \times (-\frac{2}{3}) + \frac{1}{3} \times \frac{2+\frac{2}{3}}{2} + \frac{1}{3} \times 6 = 2\frac{2}{9}$, and therefore $\gamma = (2\frac{2}{9},\ 2\frac{2}{9},\ 1\frac{5}{9})$. Players 1 and 2 each receive a side payment of $\frac{2}{9}$ from player 3.

[10] Player 1's and player 2's payoffs are identical and their strategies can be switched without impacting any payoff. This is an example of substitute players.

When player 3 is dropped, the game becomes

$$\begin{bmatrix} 2,2 & 0,0 \\ 0,0 & 1,1 \end{bmatrix},$$

and the value is $(2,2)$.

Example 5. This is an n-player game. Player 1 can choose up or down. If he chooses up each player receives 1. If he chooses down each player receives 0. The other players have no choices.

It is easy to see that $\delta_{1,k} = \max(0,\ k-(n-k))$. By formula (3), $\gamma_1 = \frac{1}{n}\sum_{k=1}^{n}\delta_{i,k} = \frac{1}{n}\sum_{k=1}^{n}\max(0,\ 2k-n)$ and, by symmetry and efficiency, $\gamma_i = \frac{n-\gamma_1}{n-1}$ for $i = 2,\ldots,n$.

Plugging in $n = 3$ *and* $= 4$ we see that in a three-player version $\gamma = \frac{1}{6}(8,5,5)$ and in a four-player version $\gamma = \frac{1}{6}(9,5,5,5)$, and it is straightforward to verify that as the number of players, n, becomes large, $\gamma \sim \frac{1}{4}(n,3,\ldots,3)$.

Thus the value of player 1 is approximately one-fourth of the total feasible output. In effect, each of the remaining players concedes one-fourth of their equal share to player 1.

This example highlights the power of player 1's threat to reduce everyone's payoff to zero. The greater the number of other players, the more player 1 derives from this threat. The value, γ, reflects this.

Example 6. Player 1 can decide which one of the $n-1$ other players will receive a contract worth 1.

Clearly, $d(1) = 0$ while $d(1 \cup S) = 1$ for any non-empty subset, S, of $N\backslash 1$. Thus $\delta_{1,1} = 0$ and $\delta_{1,k} = 1$ for $k = 2,\ldots,n$, and therefore $\gamma_1 = \frac{1}{n}\sum_{k=1}^{n}\delta_{i,k} = \frac{n-1}{n}$.

Thus, the authority to choose the winner of a contract is worth $\frac{n-1}{n}$ of the value of the contract.

Example 7. This is a variant of Example 5. Now there is a further cost, $c > 0$, to the spoiler: if player 1 chooses up then each player receives 1; if player 1 chooses down then each of the other players receives 0 but player 1 receives $-c$.

Consider the 4-player version. If $c \geq 2$ then $\gamma = (1,1,1,1)$; but if $c \leq 2$ then $\gamma_1 = \frac{1}{4}\times(-c) + \frac{1}{4}\times 0 + \frac{1}{4}\times 2 + \frac{1}{4}\times 4 = \frac{3}{2} - \frac{c}{4}$; thus $\gamma = (\frac{3}{2}-\frac{c}{4}, \frac{5}{6}+\frac{c}{12}, \frac{5}{6}+\frac{c}{12}, \frac{5}{6}+\frac{c}{12})$.

Note that the value of player 1 in this game is $1.5 - \frac{c}{4}$. Thus, the economic worth of the ability to spoil is diminished when the cost to the spoiler is increased.

It is straightforward to verify that as the number of players, n, becomes large, the value of player 1 becomes approximately one-fourth of the total payoff, the same as in Example 5. However, if the cost to the spoiler is proportional to the damage imposed on the others, say $c = c_0 n$, where $c_0 < 1$, then as the number of players becomes large, the value of player 1 becomes approximately $\frac{(1-c_0)^2}{4}$ of the total feasible output.

Example 8. This is a variant of Example 5, where there is more than one player whose approval is required for all players to receive 1. If one of the distinguished players disapproves then all players receive zero.

It is easy to compute the asymptotic behavior as $n \to \infty$. In the case of two distinguished players, the payoff to each of them divided by n – the total feasible output – converges, as $n \to \infty$, to $\int_{1/2}^{1}(2x-1)xdx = \frac{5}{24}$, which is about 21%. Thus the two spoilers receive 42% of the total feasible output, compared with 25% in the case of a single spoiler.

In the case of k distinguished players the payoff to each of them divided by n converges, as $n \to \infty$, to $\int_{1/2}^{1}(2x-1)x^{k-1}dx$. Since $k\int_{1/2}^{1}(2x-1)x^{k-1}dx$ converges, as $k \to \infty$, to 1, we see that when there are many spoilers – each with the power to reduce everyone's payoff to zero – essentially all of the economic output goes to them.

Example 9. This is a variant of Example 5, where there is more than one player whose approval is sufficient for all players to receive 1. If none of the distinguished players approves then all players receive zero.

Assume there are k distinguished players. The asymptotic behavior as $n \to \infty$ is as follows. The payoff to each distinguished player divided by n – the total feasible output – converges, as $n \to \infty$, to $\frac{2^{-k}}{k(k+1)}$. Thus the combined payoff to all the distinguished players is $\frac{2^{-k}}{(k+1)}$.

When $k = 1$ this amounts to $\frac{1}{4}$, as we have seen in Example 5. When $k = 2$ this amounts to $\frac{1}{12}$. Thus, when there are two distinguished players, only one of whose approvals is required, the fraction of the total value that they command is about 8.5%; this is in contrast to 42% in the case where both approvals are required, as in Example 8.

Remark 12. The examples demonstrate that games with more than two players exhibit phenomena that are not present in two-player games. These arise from a player's ability to play off some of the other players against one another.

6. Games of Threats

A *coalitional game of threats* is a pair (N, d), where

- $N = \{1, \ldots, n\}$ is a finite set of players.
- $d\colon 2^N \to \mathbb{R}$ is a function such that $d(S) = -d(N\backslash S)$ for all $S \subseteq N$.

Remark 13. A game of threats need not be a coalitional game as $d(\emptyset) = -d(N)$ may be non-zero.

Remark 14. If d is a game of threats then so is $-d$.

Denote by $\mathbb{D}(N)$ the set of all coalitional games of threats.

Let $\psi\colon \mathbb{D}(N) \to \mathbb{R}^n$. This may be viewed as a map that associates with any game of threats an allocation of payoffs to the players. Following Shapley [14] we consider the following axioms.

For all games of threats $(N, d_1), (N, d_2)$, and for all players i, j,

- *Symmetry* $\psi_i(d) = \psi_j(d)$ if i and j are substitutes in d (i.e., if $d(S \cup i) = d(S \cup j)$ $\forall S \subseteq N \backslash \{i, j\}$).
- *Null player* $\psi_i d = 0$ if i is a null player in d (i.e., if $d(S \cup i) = d(S)$ $\forall S \subseteq N$).
- *Efficiency* $\sum_{i \in N} \psi_i d = d(N)$.
- *Additivity* $\psi(d_1 + d_2) = \psi d_1 + \psi d_2$.

Below are two results from [7] that will be needed in the sequel.

Proposition 1. *There exists a unique map $\psi\colon \mathbb{D}(N) \to \mathbb{R}^n$ satisfying the axioms of symmetry, null player, efficiency, and additivity. It may be described as follows:*

$$\psi_i d = \frac{1}{n} \sum_{k=1}^{n} d_{i,k}, \tag{5}$$

where $d_{i,k}$ denotes the average of the $d(S)$ over all k-player coalitions that include i.

We refer to this map as the *Shapley value for games of threats.*

Definition 2. *Let $T \subseteq N$, $T \neq \emptyset$. The unanimity game of threats, $u_T \in \mathbb{D}(N)$, is defined by*

$$u_T(S) = \begin{cases} |T| & \text{if } S \supseteq T, \\ -|T| & \text{if } S \subseteq N \backslash T, \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 2. *Every game of threats is a linear combination of the unanimity games of threats u_T.*

7. Rephrasing the Main Result

Using the notion of games of threats we can provide an alternative definition of the value:

Proposition 3. *The value of a strategic game G is the Shapley value of the game of threats associated with G, i.e., $\gamma = \psi \circ \delta$, where $\gamma\colon \mathbb{G}(N) \to \mathbb{R}^n$, $\psi\colon \mathbb{D}(N) \to \mathbb{R}^n$, and $\delta\colon \mathbb{G}(N) \to \mathbb{D}(N)$ are as in (3), (5), and (2), respectively.*

Proof. Formula (3) is the same as formula (5), applied to the game of threats $d = \delta G$. □

Thus, Theorem 1 can be rephrased as follows: $\gamma = \psi \circ \delta$ *is the unique map from* $\mathbb{G}(N)$ *to* $\mathbb{R}^n$ *that satisfies the axioms of symmetry, null player, efficiency, balanced threats, and additivity.*

8. Preliminary results

In this section we present properties of the mapping $\delta\colon \mathbb{G}(N) \to \mathbb{D}(N)$ that are needed for the proof of the main result.

Let $G \in \mathbb{G}(N)$. For any $S \subseteq N$, let $\delta G(S)$ be as in (2).

Lemma 1. *δG is a game of threats.*

Proof. By the minmax theorem, $\delta G(S) = -\delta G(N\backslash S)$ for any $S \subseteq N$. □

We refer to δG as the game of threats associated with G.

Lemma 2. *$\delta : \mathbb{G}(N) \to \mathbb{D}(N)$ satisfies:*

- *$\delta(G_1 \oplus G_2) = \delta G_1 + \delta G_2$ for any $G_1, G_2 \in \mathbb{G}(N)$.*
- *$\delta(\alpha G) = \alpha\delta G$ for any $G \in \mathbb{G}(N)$ and $\alpha \geq 0$.*

Proof. Let $\mathrm{val}(G)$ denote the minmax value of the two-person zero-sum strategic game G. Then $\mathrm{val}(G_1 \oplus G_2) = \mathrm{val}(G_1) + \mathrm{val}(G_2)$.

To see this, note that by playing an optimal strategy in G_1 as well as an optimal strategy in G_2, each player guarantees the payoff $\mathrm{val}(G_1) + \mathrm{val}(G_2)$.

Now apply the above to all two-person zero-sum games played between a coalition S and its complement $N\backslash S$, as indicated in (2). □

The next lemma is an immediate consequence of the definition of δ.

Lemma 3. *$\delta : \mathbb{G}(\mathbb{N}) \to \mathbb{D}(\mathbb{N})$ satisfies:*

- *$\delta G(N) = \max_{a\in A^N}(\sum_{i\in N} g^i(a))$.*
- *If i and j are substitutes in G then i and j are substitutes in δG.*
- *If i is a null player in G then i is a null player in δG.*

Denote by $1_T \in \mathbb{R}^n$ the indicator vector of a subset $T \subseteq N$, i.e., $(1_T)_i = 1$ or 0 according to whether $i \in T$ or $i \notin T$.

Definition 3. *Let $T \subseteq N$, $T \neq \emptyset$. The unanimity strategic game on T, henceforth the unanimity game on T, is $U_T = (N, A, g_T)$, where*

$A^i = \{0, 1\}$ for all $i \in N$,

$g_T(a) = 1_T$ if $a^i = 1$ for all $i \in T$, and $g_T(a) = 0$ otherwise.

That is, if all the members of T consent then they each receive 1; however, if even one member dissents, then all receive zero; the players outside T always receive zero.

Lemma 4. *Let $T \neq \emptyset$, and let $U_T \in \mathbb{G}(N)$ be the unanimity game on T and $u_T \in \mathbb{D}(N)$ be the unanimity game of threats on T. Then $\delta U_T = u_T$.*

Proof. Consider the two-person zero-sum game between S and $N \backslash S$.

If $S \cap T \neq \emptyset, T$ then both S and $N \backslash S$ include a player in T. If these players dissent then all players receive 0. Thus the minmax value, $\delta U_T(S)$, is 0.

If $S \cap T = T$ then, by consenting, the players in S can guarantee a payoff of 1 to each player in T and 0 to all the others. Thus $\delta U_T(S) = |T|$.

If $S \cap T = \emptyset$ then, by consenting, the players in $N \backslash S$ can guarantee a payoff of 1 to each player in $T \subset N \backslash S$ and 0 to all the others. Thus $\delta U_T(S) = -|T|$.

By definition 2 , $\delta U_T = u_T$. □

Definition 4. *The anti-unanimity game on T is $V_T = (N, A, g)$, where*
$A^i = \{S \subseteq T : \ S \neq \emptyset\}$,
$g(S_1, \ldots, S_n) = \sum_{i \in T} -1_{S_i}$.

That is, each player in T chooses a non-empty subset of T where each member loses 1. Players outside T also choose such subsets, but their choices have no impact. Thus the payoff to any player, i, is minus the number of players in T whose chosen set includes i.

Lemma 5. $\delta V_T = -u_T$.

Proof. Let S be a subset of N such that $T \subseteq S$. In the zero-sum game between S and its complement, each player in S chooses a subset of T of size 1. Thus $\delta V_T(S) = -|T|$.

Let S be a subset of N such that $T \cap S \neq \emptyset$ and $T \setminus S \neq \emptyset$. In the zero-sum game between S and its complement, the minmax strategies are for the players in S to choose $T \setminus S$ and for the players in $N \backslash S$ to choose $T \cap S$. The resulting payoff is $-t_1 t_2 - (-t_2 t_1) = 0$, where t_1 and t_2 are the number of elements of $T \cap S$ and $T \setminus S$, respectively. Thus $\delta V_T(S) = 0$.

Therefore, $\delta V_T = -u_T$. □

Lemma 6. *For every game of threats $d \in \mathbb{D}(N)$ there exists a strategic game $U \in \mathbb{G}(N)$ such that $\delta U = d$. Moreover, there exists such a game that can be expressed as a direct sum of non-negative multiples of the unanimity games $\{U_T\}_{T \subseteq N}$ and the anti-unanimity games $\{V_T\}_{T \subseteq N}$.*

Proof. By Proposition 2, d is a linear combination of the unanimity games of threats u_T.

$$d = \sum_T \alpha_T u_T - \sum_T \beta_T u_T \text{ where } \ \alpha_T, \beta_T \geq 0 \text{ for all } T.$$

By Lemmas 4 and 5,

$$d = \sum_T \delta(\alpha_T U_T) + \sum_T \delta(\beta_T V_T),$$

and, by Lemma 2,

$$d = \delta((\bigoplus_{T \subseteq N} \alpha_T U_T) \oplus (\bigoplus_{T \subseteq N} \beta_T V_T)),$$

where $\bigoplus_T$ stands for the direct sum of the games parameterized by T. □

Remark 15. In particular, Lemma 6 establishes that the mapping $\delta\colon \mathbb{G}(N) \to \mathbb{D}(N)$ is onto.

As was pointed out earlier, the operation $\oplus$ does not have a natural inverse. However, we have the following:

Lemma 7. *For every $G \in \mathbb{G}(N)$ there exists a $\delta-$inverse, i.e., $U \in \mathbb{G}(N)$ such that $\delta(G \oplus U) = 0$. Moreover, if $G' \in \mathbb{G}(N)$ is such that $\delta G' = \delta G$ then there exists $U \in \mathbb{G}(N)$ that is a $\delta-$inverse of both G and G'.*

Proof. Consider $-\delta G \in \mathbb{D}(N)$. By Lemma 6, there exists $U \in \mathbb{G}(N)$ such that $-\delta G = \delta U$. By Lemma 2, $\delta(G \oplus U) = 0$.

And if G' is such that $\delta G' = \delta G$ then, by the same argument, $\delta(G' \oplus U) = 0$. □

Lemma 8. *Assume that $\gamma\colon \mathbb{G}(N) \to \mathbb{R}^n$ satisfies the axioms of balanced threats and efficiency. If $\delta G = 0$ then $\gamma G = 0$.*

Proof. Since $(\delta G)(S) = 0$ for all subsets of N, the axiom of balanced threats implies that all the $\gamma_i G$ are the same. By efficiency, their sum is equal to $\max_{a \in A^N}(\sum_{i \in N} g^i(a)) = \delta G(N) = 0$. Thus each of the $\gamma_i G$ is zero. □

Proposition 4. *If $\gamma\colon \mathbb{G}(N) \to \mathbb{R}^n$ satisfies the axioms of balanced threats, efficiency, and additivity then γG is a function of δG.*

Proof. Let $G, G' \in \mathbb{G}(N)$ be such that $\delta G = \delta G'$. We must show that $\gamma G = \gamma G'$.

By Lemma 7, there exists $U \in \mathbb{G}(N)$ such that $\delta(G \oplus U) = 0 = \delta(G' \oplus U)$.

By Lemma 8, $\gamma(G \oplus U) = 0 = \gamma(G' \oplus U)$.

Thus, by the additivity axiom, $\gamma G = -\gamma U = \gamma G'$. □

Lemma 9. *For any $T \neq \emptyset$ and $\alpha \geq 0$, the axioms of symmetry, null player, and efficiency determine γ on the game αU_T. Specifically, $\gamma(\alpha U_T) = \alpha 1_T$.*

Proof. Any $i \notin T$ is a null player in U_T, and so $\gamma_i = 0$. Any $i, j \in T$ are substitutes in U_T, and so $\gamma_i = \gamma_j$. By efficiency, the sum of the γ_i is the maximum total payoff, which, since $\alpha > 0$, is $\alpha|T|$. Thus each of the $|T|$ non-zero γ_i is equal to α. □

Lemma 10. *For any $\alpha \geq 0$, the axioms (of symmetry, null player, additivity, balanced threats, and efficiency) determine γ on the game αV_T. Specifically, $\gamma(\alpha V_T) = -\alpha 1_T$.*

Proof. By Lemma 9 the axioms determine $\gamma(\alpha U_T) = \alpha 1_T$. By Lemmas 4 and 5, $\delta(\alpha V_T \oplus \alpha U_T) = 0$. Therefore, by Lemma 8, $\gamma(\alpha V_T \oplus \alpha U_N) = 0$. Thus, by additivity, $\gamma(\alpha V_T) = -\gamma(\alpha U_T) = -\alpha 1_T$. □

Remark 16. We cannot rely on the same proof as that of Lemma 9, by appealing to symmetry and efficiency. In the game V_T, it is not true that any two players, $i, j \in T$, are substitutes, because the payoff functions are not identical. If we had adopted a more restrictive version of the symmetry axiom – that the names of the players don't matter – then any $i, j \in T$ would be substitutes and the direct proof would be valid. But this more restrictive version of the axiom would lead to a weaker uniqueness theorem.

Proposition 5. *The map γ of formula (3) satisfies the axiom of individual rationality.*

Proof. Let $G = (N, A, g)$ be a strategic game. By symmetry, it is sufficient to prove individual rationality for player 1, i.e., that $\gamma_1 G \geq \pi^1$, where π^1 denotes player 1's security level.

Let S_1, S_2 be a partition of $N\backslash 1$. We claim that

$$(\delta G)(S_1 \cup 1) + (\delta G)(S_2 \cup 1) \geq 2\pi^1. \tag{6}$$

To see this, let $\bar{x}^1$ be a strategy that guarantees player 1 his security level, i.e.,

$$\min_{x^{N\backslash 1} \in X^{N\backslash 1}} g^1(\bar{x}, x^{N\backslash 1}) = \max_{x^1 \in X^1} \min_{x^{N\backslash 1} \in X^{N\backslash 1}} g^1(x^1, x^{N\backslash 1}) = \pi^1. \tag{7}$$

We have:

$$\begin{aligned}
& \delta G(S_1 \cup 1) \\
= \quad & \max_{x \in X^{S_1 \cup 1}} \min_{y \in X^{S_2}} \left(\sum_{i \in S_1 \cup 1} g^i(x, y) - \sum_{i \in S_2} g^i(x, y) \right) \\
\geq \quad & \max_{x \in X^{S_1}} \min_{y \in X^{S_2}} \left(\sum_{i \in S_1 \cup 1} g^i(\bar{x}^1, x, y) - \sum_{i \in S_2} g^i(\bar{x}^1, x, y) \right) \\
\geq \quad & \pi^1 + \max_{x \in X^{S_1}} \min_{y \in X^{S_2}} \left(\sum_{i \in S_1} g^i(\bar{x}^1, x, y) - \sum_{i \in S_2} g^i(\bar{x}^1, x, y) \right). \qquad (8)
\end{aligned}$$

The first inequality follows since restricting the set of available strategies cannot increase the maximum of a function, and the second inequality follows from (7) and the fact that the maxmin of a function is monotonic in that function.

Similarly, we have

$$\delta G(S_2 \cup 1) \geq \pi^1 + \max_{x \in X^{S_2}} \min_{y \in X^{S_1}} \left(\sum_{i \in S_2} g^i(\bar{x}^1, x, y) - \sum_{i \in S_1} g^i(\bar{x}^1, x, y) \right). \tag{9}$$

By the minmax theorem, the sum of the right-hand sides of (8) and (9) is $2\pi^1$; therefore, adding these two inequalities implies (6).

Now, as S_1 ranges over all the sets of size $k-1$ that do not include 1, S_2 ranges over all the sets of size $n-k$ that do not include 1; thus $S_1 \cup 1$ ranges over all the sets of size k that include 1 and $S_2 \cup 1$ ranges over all the sets of size $n-k+1$ that include 1. Taking the average of inequality (6) over all these sets we have

$$\delta_{1,k} + \delta_{1,n-k+1} \geq 2\pi^1,$$

where $\delta_{1,k}$ denotes the average of the $\delta G(S)$ over all k-player coalitions that include 1.

Taking the average over $k = 1, \ldots, n$ we obtain

$$2 \times \frac{1}{n} \sum_{k=1}^{n} \delta_{1,k} \geq 2\pi^1.$$

Thus, by formula (3), $\gamma_1 G \geq \pi^1$.

□

9. Proof of the main result

.

Proof. of Theorem 1.

We first prove uniqueness.

Let $G \in \mathbb{G}(N)$. Consider $\delta G \in \mathbb{D}(N)$; by Lemma 6 there exists a game $U \in \mathbb{G}(N)$ that is a direct sum of non-negative multiples of the unanimity games $\{U_T\}_{T \subseteq N}$ and the anti-unanimity games $\{V_T\}_{T \subseteq N}$, such that $\delta G = \delta U$.

By Proposition 4, $\gamma G = \gamma U$ and so it suffices to show that γU is determined by the axioms.

Now, by Lemmas 9 and 10 , γ is determined on non-negative multiples of the unanimity games $\{U_T\}_{T \subseteq N}$ and the anti-unanimity games $\{V_T\}_{T \subseteq N}$. It then follows from the axiom of additivity that γ is determined on U.

To prove existence we show that the value, $\gamma = \psi \circ \delta$, satisfies the axioms.

Efficiency, symmetry, and the null player axiom follow from Lemma 3 and the corresponding properties of the Shapley value ψ.

Additivity follows from Lemma 2 and the linearity of the Shapley value.

To see that γ satisfies the axiom of balanced threats, assume that $(\delta G)S = 0$ for any proper subset of N. Then $\delta_{i,k}$, the average of the $(\delta G)(S)$ over all k-player coalitions that include i, is zero for any $k < n$. It then follows from (3) that $\gamma_i G = \frac{1}{n}\delta G(N)$; thus $\gamma_i G = \gamma_j G$ for all i, j.

Finally, Proposition 5 establishes that γ satisfies the axiom of individual rationality. □

10. The Shapley value of strategic games

Recall that the (von Neumann–Morgenstern–) Shapley value of a strategic game G is the Shapley value of the coalitional game vG that is defined by

$$(vG)(S) := \max_{x\in X^S} \min_{y\in X^{N\setminus S}} \sum_{i\in S} g^i(x,y). \tag{10}$$

In Section 2 we noted that this concept, which we referred to as the the VNM–Shapley value, is identical to the value in constant-sum games, but we argued that it is less convincing in general-sum games. Here we provide a characterization that parallels the characterization of the value and clarifies the relationship between the two concepts.

Let $G \in \mathbb{G}(N)$. Define

$$(\hat{\delta}G)(S) := (vG)(S) - (vG)(N\setminus S). \tag{11}$$

We introduce the following axiom. For all $G \in \mathbb{G}(N)$

Balanced Security Levels If for every subset S, $v(S) = v(N\setminus S)$ then $\hat{\gamma}_i G = \hat{\gamma}_j G$ for all $i, j \in N$.

An alternative formulation, which highlights the similarity with the axiom of balanced threats, is as follows. If $(\hat{\delta}G)(S) = 0$ for every subset S, then $\hat{\gamma}_i G = \hat{\gamma}_j G$ for all $i, j \in N$.

Proposition 6. *The VNM–Shapley value is the unique map from $\mathbb{G}(N)$ to $\mathbb{R}^n$ that satisfies the axioms of symmetry, null player, efficiency, balanced security levels, and additivity. It may be described as follows:*

$$\hat{\gamma}_i G = \frac{1}{n}\sum_{k=1}^{n} \hat{\delta}_{i,k}, \tag{12}$$

where $\hat{\delta}_{i,k}$ denotes the average of the $(\hat{\delta}G)(S)$ over all k-player coalitions that include i. Furthermore, this map satisfies the axiom of individual rationality.

Uniqueness can be proved in the same way as in Theorem 1. It is straightforward to verify that all the lemmas that involve δ remain valid when δ is replaced by $\hat{\delta}$. In particular, note that $\hat{\delta}U_T = \delta U_T = u_T$ and $\hat{\delta}V_T = \delta V_T = -u_T$.

Recall that ψ denotes the Shapley value for games of threats. The proof that $\psi \circ \hat{\delta}$ satisfies the axioms is similar to the proof in Theorem 1 that $\psi \circ \delta$ satisfies the axioms of that theorem. The proof that $\hat{\gamma} = \psi \circ \hat{\delta}$ is similar to the proof of Proposition 3.

11. n-player Bargaining with variable threats

Harsanyi's [5] original definition of the value for a strategic game G is based on a generalization of Nash's [11] two-person model of bargaining with variable threats. In an initial stage, each coalition S declares a correlated strategy x^S as its threat. Define a payoff $w_x(S) := \sum_{i\in S} g^i(x^S, x^{N\setminus S})$ to the coalition S and a payoff $w_x(N\setminus S) :=$

$\sum_{i\in N\setminus S} g^i(x^S, x^{N\setminus S})$ to $N\setminus S$. Given the threat strategies of all the coalitions, and the resulting coalitional game w_x, each coalition receives a "settlement," defined as the Shapley value of w_x.

Harsanyi considers an auxiliary $(2^n - 1)$-player strategic game where the players are the non-empty coalitions S, where each player chooses a threat strategy x^S, and where the payoff to player S is the sum of the payoffs to the members of S in the Shapley value of the coalitional game w_x. (Recall that w_x depends on the threats of all the coalitions.) Harsanyi then shows that, while this game may have many Nash equilibrium points, they all have the same payoffs. These payoffs are taken as the definition of the value of the original n-person strategic game G.

Harsanyi notes that, for any coalitional game w, the Shapley value of player i is a weighted average of the differences $w(S) - w(N \setminus S)$ over the coalitions S that include i. Thus, in any Nash equilibrium of the $(2^n - 1)$-player auxiliary game, player S strives to maximize $w_x(S) - w_x(N \setminus S)$ while player $N\setminus S$ strives to minimize the same expression; therefore the coalitional game w_x that results from any Nash equilibrium $(x^S)_{S\subset N}$ is such that $w_x(S) - w_x(N\setminus S) = \max_{x\in X^S} \min_{y\in X^{N\setminus S}} \left(\sum_{i\in S} g^i(x,y) - \sum_{i\notin S} g^i(x,y)\right) = (\delta G)(S)$. It is then straightforward to verify that the Shapley value of the coalitional game w_x is equal to the Shapley value of the game of threats δG. This establishes that the above-described scheme indeed gives rise to the concept that we call "the value."

Myerson [10] takes issue with the use of the Shapley value as the settlement function in Harsanyi's scheme. A major concern expressed by Myerson is that the value does not satisfy a notion of individual rationality, which he defines as follows. For each game G and a vector of equilibrium threats x of the coalitions, $\gamma_i G$ is at least the maximum of $g_i(a, x(N \setminus i))$, where a ranges over all the strategies A^i of player i. He proposes to replace the Shapley value by another settlement function: the Shapley value in partition function form, which he characterizes as the unique settlement function that satisfies – as a settlement function – two desirable properties.

The approach presented in this paper takes a different route to generalizing Nash's "bargaining with variable threats" to n-player games. Instead of trying to generalize the process, we attempt to generalize a key property. In a two-player game, if the minmax value of the game of differences is zero, then the Nash variable-threats solution allocates the same payoff to both players. This can be interpreted as saying that if neither player has threat power, then each receives the same payoff. Now, in a two-person game the only proper subsets are the two singletons, therefore the only threats to consider are those by one player aganist the other. But in n-person games there are many proper subsets, each of which can threaten its complement. If we wish to generalize the Nash variable-threats solution to n-person games, then it seems reasonable to require that in games where no proper coalition has threat power, i.e., $d(S) = 0$ for all proper subsets of N, all players receive the same payoff. This is a less restrictive variant of our axiom of balanced threats, which requires in addition that the maximal sum of the payoffs be zero; see Remark 3.

12. The coco value

Kalai and Kalai [6] introduced the "coco value," which coincides with the value of two-person strategic games. Their main result is an axiomatic characterization of the coco value.

The axioms that they consider are the following: efficiency, shift invariance (if G_α is a modification of G obtained by adding to the payoff of one player, say player 1, an amount α everywhere, then $\gamma G_\alpha = \gamma G + (\alpha, 0)$), invariance to redundant strategies (removing a duplicate row or column in the payoff matrix does not affect the value), monotonicity in actions (removing a pure strategy of a player cannot increase the player's value), and payoff dominance (if player 1's payoff is everywhere strictly greater than player 2's payoff then $\gamma_1 \geq \gamma_2$). They prove that there is a unique map from $\mathbb{G}(2)$ to $\mathbb{R}^2$ that satisfies these axioms.

Since Kalai and Kalai characterize the same concept for two-person games as we do, their axioms are equivalent to ours in the two-person case. To see a direct connection between the two sets of axioms, it may be helpful to note that in two-person games the general additivity axiom can be replaced by the requirement that the solution be additive over the direct sum of a game and a trivial game, which amounts to shift invariance.

In games with more than two players the value still satisfies all the Kalai and Kalai axioms other than payoff dominance. This follows from Remark 5 and Proposition 10. However, the value does not satisfy payoff dominance. This is a reflection of the more complex considerations in games with more than two players. In Example 7, player 1's payoff is everywhere smaller than player 2's, but 1's value is greater. This is so because of player 1's ability to play off some of his opponents against each other.

13. Appendix A. Additional Properties of the value

The following is another desirable property of a value solution.

Small worlds If the set of players is the union of two disjoint subsets such that the payoffs to the players in each subset are unaffected by the actions of the players in the other subset, then the value of each player is the same as it would be in the game restricted to the subset that includes the player.

Proposition 7. *The value satisfies the small-worlds axiom.*

Proof. Let $G = (N, A, g)$, where $N = N_1 \cup N_2$, $N_1 \cap N_2 = \emptyset$, and where the actions of players in N_1 do not affect the payoffs to players in N_2, and vice versa.

Assume, w.l.o.g., that $1 \in A_2^i$ for all $i \in N$. Define $G_1 \in \mathbb{G}(N)$ by modifying G as follows. Restrict the set of pure strategies of each player in N_2 to $\{1\}$ and define $g_1^i = g^i$ for $i \in N_1$ and $g_1^i = 0$ for $i \in N_2$; and define G_2 in a similar way.

By the definition (2) of δ,

$$\delta G = \delta G_1 + \delta G_2.$$

Recall that $\gamma = \psi \circ \delta$, where ψ is the Shapley value for games of threats (Proposition 3). Since ψ is additive,

$$\gamma G = \psi \circ \delta(G) = \psi \circ \delta(G_1 + G_2) = \psi \circ \delta G_1 + \psi \circ \delta G_2 = \gamma G_1 + \gamma G_2.$$

Since any $i \in N_1$ is a null player in G_2, it follows from the null-player axiom that $\gamma_i G_2 = 0$. Thus

$$\gamma_i G = \gamma_i G_1 + \gamma_i G_2 = \gamma_i G_1 \text{ for all } i \in N_1.$$

Similarly, $\gamma_i G = \gamma_i G_2$ for all $i \in N_2$.

Thus, for $i \in N_1$, the value of G is the same as the value of G_1, which may be viewed as the restriction of G to N_1; and similarly for $i \in N_2$. □

Remark 17. It is insufficient to assume that the payoffs to the players in N_1 are unaffected by the actions of the players in N_2. In Example 4, player 3 has only one strategy and so he obviously cannot affect the payoffs of the other players. Yet when player 3 is dropped, the values for players 1 and 2 change.

Remark 18. The small-worlds axiom may be viewed as an instance of the more general statement, that the additivity of the value extends to games over two different sets of players. Let $G_1 \in \mathbb{G}(N_1)$ and $G_2 \in \mathbb{G}(N_2)$. By adding the members of $N_2 \setminus N_1$ as dummy players in G_1, and the members of $N_1 \setminus N_2$ as dummy players in G_2, we may view both G_1 and G_2 as games in $\mathbb{G}(N_1 \cup N_2)$. Thus $\gamma(G_1 \oplus G_2) = \gamma G_1 + \gamma G_2$. Since the value of the existing players is unaffected by the addition of dummy players, $\gamma_i(G_1 \oplus G_2) = \gamma_i G_1$ for all $i \in N_1 \setminus N_2$ and $\gamma_i(G_1 \oplus G_2) = \gamma_i G_2$ for all $i \in N_2 \setminus N_1$. The small-worlds axiom corresponds to the case where N_1 and N_2 are disjoint.

Recall that the axiom of balanced threats says that if $\delta G(S) = 0$ for any subset S, then $\gamma_i G = \gamma_j G$ for all i, j.

We now consider a stronger version of this axiom.

Strong axiom of balanced threats If $\delta G(S) = 0$ for all subsets S such that $S \ni i$ and $S \not\ni j$ then $\gamma_i G = \gamma_j G$.

Proposition 8. *The value satisfies the strong axiom of balanced threats.*

Proof. Since $(\delta G)S = 0$ for subsets S that include i but do not include j, formula (4) becomes

$$\gamma_i G = \frac{1}{n} \sum_{k=1}^{n} \frac{1}{\binom{n-1}{k-1}} \sum_{\substack{S:\{i,j\} \subseteq S \\ |S|=k}} (\delta G)(S).$$

But the r.h.s. is the same for $\gamma_j G$. □

Recall that our axiom of symmetry says that if two players, i and j, are substitutes in the game G then $\gamma_i G = \gamma_j G$. By contrast, the classic axiom requires that the value be invariant to permutations of the players' names. Since every permutation of N consists of a sequence of pairwise exchanges, the axiom can be stated as follows.

Axiom of full symmetry Let $G = (N, A, g)$ and let $\hat{G} = (N, \hat{A}, \hat{g})$ be such that $\hat{A}^i = A^j$, $\hat{A}^j = A^i$ and $\hat{g}^i = g^j$, $\hat{g}^j = g^i$; then $\gamma_i \hat{G} = \gamma_j G$, $\gamma_j \hat{G} = \gamma_i G$, and $\gamma_k \hat{G} = \gamma_k G$ for $k \neq i, j$.

Clearly, this axiom is stronger than our symmetry axiom. Still, formula (3) establishes that

Proposition 9. *The value satisfies the axiom of full symmetry.*

Given a game $G = (N, A, g)$ and $\alpha \in \mathbb{R}^n$, let $G + \alpha$ be the game obtained from G by adding α to each payoff entry, namely, $G + \alpha = (N, A, g + \alpha)$.

Axiom of shift invariance $\gamma(G + \alpha) = \gamma G + \alpha$.

Proposition 10. *The value satisfies the axiom of shift invariance.*

Proof. The definition (2) of δ implies that

$$\delta(G + \alpha)(S) = (\delta G)(S) + \sum_{j \in S} \alpha_j - \sum_{j \in N \setminus S} \alpha_j.$$

Therefore, if $i \in S$ then $\delta(G + \alpha)(S) + \delta(G + \alpha)(i \cup (N \setminus S)) = (\delta G)(S) + \sum_{j \in S} \alpha_j - \sum_{j \in N \setminus S} \alpha_j + (\delta G)(i \cup (N \setminus S)) + \sum_{j \in i \cup N \setminus S} \alpha_j - \sum_{i \neq j \in S} \alpha_j = (\delta G)(S) + (\delta G)(i \cup (N \setminus S)) + 2\alpha_i$. As the map from subsets S of size k that contain player i, defined by $S \mapsto i \cup (N \setminus S)$, is 1-1 and onto the subsets of size $n - k + 1$ that contain player i, we deduce that $\delta_{i,k}(G + \alpha) + \delta_{i,n-k+1}(G + \alpha) = \delta_{i,k}(G) + \delta_{i,n-k+1}(G) + 2\alpha_i$. Therefore, by formula (3) for the value, $\gamma(G + \alpha) = \gamma G + \alpha$. ☐

Remark 19. Let I_α be a game where the payoff is the constant α. The game $G + \alpha$ is strategically equivalent to $G \oplus I_\alpha$. As the valuesof two strategically equivalent games coincide, it would have been sufficient to prove that $\gamma(G \oplus I_\alpha) = \gamma G + \alpha$. For this equality one need not rely on the axiom of balanced threats.

Proposition 11. *A map $\gamma : \mathbb{G}(N) \to \mathbb{R}^n$ that satisfies the additivity, efficiency, and null player axioms satisfies $\gamma(G \oplus I_\alpha) = \gamma G + \alpha$.*

Proof. We prove that $\gamma I_\alpha = \alpha$.

Note that a player i is a null player in I_α if and only if $\alpha_i = 0$. If all the players in I_α are null players then $\alpha = 0$ and $\gamma(I_\alpha) = 0$ by the null-player axiom. Assume that there is one non-null player in I_α, say player i. Then, $\alpha_i \neq 0$ and $\forall j \neq i$, $\alpha_j = 0$, and by the null-player axiom $\gamma_j(I_\alpha) = 0$, and by the efficiency axiom $\gamma_i(I_\alpha) = \alpha_i$. Therefore, $\gamma(I_\alpha) = \alpha$.

We continue by induction on the number of non-null players in I_α. If there are $k > 1$ non-null players in I_α, then let $\alpha(1)$ and $\alpha(2)$ be such that $\alpha = \alpha(1) + \alpha(2)$ and in each game $I_{\alpha(1)}$ and $I_{-\alpha(2)}$ there are fewer than k non-null players. By the additivity axiom,

$\gamma(I_\alpha \oplus I_{-\alpha(2)}) = \gamma(I_\alpha) + \gamma(I_{-\alpha(2)})$, and by the induction hypothesis $\gamma(I_\alpha \oplus I_{-\alpha(2)}) = \alpha(1)$ and $\gamma(I_{-\alpha(2)}) = -\alpha(2)$. We conclude that $\gamma(I_\alpha) = \alpha(1) + \alpha(2) = \alpha$.

Therefore, $\gamma(G \oplus I_\alpha) = \gamma G + \gamma I_\alpha = \gamma G + \alpha$, where the first equality follows from the axiom of additivity and the second equality from the previously proved $\gamma I_\alpha = \alpha$. □

We end this section by stating a property of the value that we mentioned in Section 2.

Proposition 12. *The value coincides with the VNM–Shapley value in constant-sum games and in games without externalities, where the sum of the payoffs to the players in any coalition depends only on the strategies of the players belonging to that coalition.*

Proof. If the strategic game is constant-sum, then an optimal strategy for S in the problem $\max_{x \in X^S} \min_{y \in X^{N \setminus S}} \sum_{i \in S} g^i(x, y)$ is also an optimal strategy for S in the problem $\max_{y \in X^{N \setminus S}} \min_{x \in X^S} \sum_{i \in N \setminus S} g^i(x, y)$. This is also true, trivially, in a game without externalities, where any strategy for S is optimal in minimizing the total payoff to $N \setminus S$. In both these cases, then, the minmax strategies in the two person zero-sum game where the payoff (to player 1) is the total payoff to S, are also minmax strategies in the game where the payoff is the difference between the total payoff to S and the total payoff to $N \setminus S$. Thus the optimal values in (10) and in (2) are the same. □

An important example of a strategic game without externalities is an exchange economy. The proposition shows that the value and the VNM–Shapley value coincide there. However, the two notions are different in exchange economies with majority-determined taxation (Aumann, Kurtz, and Neyman [3].)

14. Appendix B. The axioms for the value are tight

In this section we show that the axioms for the value are tight; i.e., if any one of them is dropped then the uniqueness theorem is no longer valid. Furthermore, the axioms are tight even if balanced threats and symmetry are replaced by their more restrictive versions (strong balanced threats and full symmetry, respectively).

Let, for all $i \in N$,

$$\gamma_i G = \frac{1}{n} (\delta G)(N). \tag{13}$$

It is easy to verify that

Claim 1. *The mapping $\gamma \colon \mathbb{G}(N) \to \mathbb{R}^n$ defined by (13) satisfies all the axioms except for the null-player axiom.*

Let, for all $i \in N$,

$$\gamma_i G = 0. \tag{14}$$

It is easy to verify that

Claim 2. *The mapping $\gamma\colon \mathbb{G}(N) \to \mathbb{R}^n$ defined by (14) satisfies all the axioms except for efficiency.*

For each integer $1 \le k \le n$, let π_k be the order $k, k+1, \ldots, n, 1, \ldots, k-1$, and let, for all $i \in N$,

$$\gamma_i G = \frac{1}{2n} \sum_{k=1}^{n} (\delta G(\mathcal{P}_i^{\pi_k} \cup i) - \delta G(\mathcal{P}_i^{\pi_k})), \tag{15}$$

where $P_i^{\pi_k}$ consists of all players j that precede i in the order π_k.

Claim 3. *The mapping $\gamma\colon \mathbb{G}(N) \to \mathbb{R}^n$ defined in (15) satisfies all the axioms except for symmetry.*

Proof. It is easy to verify that the axioms of null player, balanced threats, and additivity are satisfied. As for efficiency, it is sufficient to verify it for G such that δG is a unanimity game in $\mathbb{D}(N)$.

Let then δG be the unanimity game on T, i.e., $\delta G(S) = 1$ if $S \supseteq T$, -1 if $S \subseteq N \backslash T$, and zero otherwise.

For $i \in T$, $\delta G(\mathcal{P}_i^{\pi_k} \cup i) = 1$ if $\mathcal{P}_i^{\pi_k} \cup i \supseteq T$, i.e., if in the order π_k, i is the last among the members of T, and zero otherwise. Thus

$$\sum_{i \in T} \frac{1}{n} \sum_{k=1}^{n} \delta G(\mathcal{P}_i^{\pi_k} \cup i) = \frac{1}{n} \sum_{k=1}^{n} \sum_{i \in T} \delta G(\mathcal{P}_i^{\pi_k} \cup i) = \frac{1}{n} \sum_{k=1}^{n} 1 = 1,$$

where the third equality follows from the fact that in each order π_k exactly one $i \in T$ is last among the members of T.

Similarly, for $i \in T$, $\delta G(\mathcal{P}_i^{\pi_k}) = -1$ if $\mathcal{P}_i^{\pi_k} \subseteq N \backslash T$, i.e., if in the order π_k, i is the first among the members of T, and zero otherwise. Since in each order π_k exactly one $i \in T$ is first among the members of T, we have

$$\sum_{i \in T} \frac{1}{n} \sum_{k=1}^{n} \delta G(\mathcal{P}_i^{\pi_k}) = \frac{1}{n} \sum_{k=1}^{n} \sum_{i \in T} \delta G(\mathcal{P}_i^{\pi_k}) = \frac{1}{n} \sum_{k=1}^{n} (-1) = -1.$$

By (15), $\sum_{i \in T} \gamma_i = \frac{1}{2}(1+1) = 1$.

For $i \notin T$, $\mathcal{P}_i^{\pi_k} \cup i \subseteq T$ if and only if $\mathcal{P}_i^{\pi_k} \subseteq T$, and $\mathcal{P}_i^{\pi_k} \cup i \subseteq N \backslash T$ if and only if $\mathcal{P}_i^{\pi_k} \subseteq N \backslash T$. By (15) then, $\gamma_i G = 0$.

Thus $\sum_{i=1}^{n} \gamma_i = 1 = \delta G(N)$, completing the proof of efficiency.

To see that γ of equation (15) does not satisfy the symmetry axiom, consider the unanimity game on $\{1, 2, 5\}$ in the game with player set $\{1, \ldots, 5\}$.

Player 1 is first in T for the order π_1 and last in T for the order π_2. Thus $\gamma_1 = \frac{1}{10}(1-(-1)) = \frac{1}{5}$.

Player 2 is first in T for the order π_2 and last in T for the orders π_3, π_4 and π_5. Thus $\gamma_2 = \frac{1}{10}(1-(-3)) = \frac{2}{5}$.

But 1 and 2 are substitutes. □

Next, observe that the VNM–Shapley value for strategic games satisfies all the axioms except for the axiom of balanced threats. (See Section 10.)

Finally, consider the following map. All dummy players in G receive the same as in the value formula(3), and the others share equally the remainder relative to $(\delta G)(N)$. It is easy to verify that this solution satisfies all the axioms except for additivity. (It does not even satisfy consensus-shift invariance.)

We conclude this section by commenting on the axioms required to imply that the value, γG, is a function of δG. The axiom of balanced threats says that if $(\delta G)(S) = 0$ for any subset S then $\gamma G = 0$. It would seem then that this axiom alone would suffice. However, this is not the case.

Let δG and vG be as defined in (2) and (10), respectively, and fix $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ s.t. $f(0,y) = f(x,0) = 0$, and $f(x,x) = x \ \ \forall x, y$.

Define $\gamma(G)$ as the Shapley value of the coalitional game u with $u(S) := f(d(G)(S), v(G)(S))$.

Claim 4. *The mapping $\gamma\colon \mathbb{G}(N) \to \mathbb{R}^n$ defined above satisfies the axioms of balanced threats, symmetry, efficiency, and null player, but it is not a function of δG.*

References

[1] Aumann, R.J. and M. Kurtz (1977), Power and Taxes, *Econometrica*, 45, 1137–1161.

[2] Aumann, R.J. and M. Kurtz (1977), Power and Taxes in a Multi-Commodity Economy, *Israel Journal of Mathematics*, 27, 185–234.

[3] Aumann, R.J., M. Kurtz, and A. Neyman (1983), Voting for Public Goods, *Review of Economic Studies*, L(4), 677–693.

[4] Aumann, R.J., M. Kurtz, and A. Neyman (1987), Power and Public Goods, *Journal of Economic Theory*, 42, 108–127.

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

[6] Kalai, A. and E. Kalai (2013), Cooperation in Strategic Games Revisited, *Quarterly Journal of Economics*, 128, 917–966.

[7] Kohlberg, E. and A. Neyman (2016), Games of Threats, , Federmann Center for the Study of Rationality discussion paper 706.

[8] Luce, R.D. and H. Raiffa (1957), *Games and Decisions*, John Wiley and Sons.

[9] Maskin, E. (2008), How Can Cooperative Game Theory be Made More Relevant to Economics? In Nash, J. and J. Rassias *Open Problems in Mathematics*, Springer International Publishing.

[10] Myerson, R. (1978), Threat Equilibria and Fair Settlements in Cooperative Games, *Mathematics of Operations Research*, 3, 265–274.

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

[12] Ray, D. and R. Vohra (1997), Equilibrium Binding Agreements, *Journal of Economic Theory*, 73, 30–78.
[13] Shapley, L. (1951), Notes on the n-Person Game – II: the Value of an n-Person Game, RAND research memorandum ATI 210720.
[14] Shapley, L. (1953), A Value for n-Person Games. In Kuhn H.W. and Tucker, A.W. (eds.), *Contributions to the Theory of Games*, Annals of Mathematics Studies, 28, 307–319.
[15] Shapley, L. (1984), Mathematics 147 Game Theory, Lecture Notes, UCLA Department of Mathematics, 1984, 1987, 1988, 1990.
[16] von Neumann, J. and O. Morgenstern (1953), *Theory of Games and Economic Behavior*, third edition, Princeton University Press.