Reevaluating the Shapley Value: A New Justification and Extension

By Adam Brandenburger and Barry Nalebuff*

Inspired by the bargaining procedure of Shapley (1953), we introduce a novel procedure in which the marginal contribution of the player joining a coalition is split in any proportion between that player and the members of the coalition being joined. Surprisingly, this more general procedure also leads to the Shapley value. Departing from the random-ordering assumption underneath the Shapley value (and our generalized procedure) leads to a new extension of the Shapley value where weights on players depend endogenously on marginal contributions.

Keywords: Shapley value, n-person bargaining, generalized bargaining procedures, value division, weighted Shapley value.

The Shapley value (Shapley, 1953) results from an axiomatic approach to the allocation of value in a cooperative game. It has an elegant procedural implementation: let the grand coalition of all players be constructed by players joining one at a time in a random order. In each ordering, give all the marginal value created by the new player joining the coalition to that player. The expected value is the Shapley value.

In this procedure, which we will call the Shapley procedure, the "negotiation" between the new player and the existing players involves an extreme division of value. The player who joins a coalition receives the full amount of the benefit created. This feature has been commented on by Brock (1992):

> A criticism of this [Shapley's] scheme is that a player is awarded his entire (utility) contribution to a coalition. This violates an intuitive symmetry whereby we would expect the player to receive only a portion of his contribution to the coalition.

While each of the possible orderings in the Shapley procedure leads to an extreme division of value which appears unfair, the average across all the unfair options may be seen as fair. The unfairness averages out. Even so, at an intuitive level, we think a procedure would be more attractive were the gains more evenly balanced among the parties in each possible ordering.

* NYU Stern School of Business, NYU Tandon School of Engineering, NYU Shanghai, New York University, adam.brandenburger@stern.nyu.edu; and School of Management, Yale University, barry.nalebuff@yale.edu. We thank seminar audiences at Washington University in St. Louis, Yale University, and Bianca Battaglia, Jack Fanning, Elio Farina, John Geanakoplos, Sergiu Hart, Roberto Serrano, Kai Hao Yang, and Jidong Zhou for very helpful comments. Financial support from NYU Stern School of Business, NYU Shanghai, and J.P. Valles is gratefully acknowledged.

In our procedure, the player joining the existing coalition shares the gains with the members of the coalition in the proportion $\alpha\colon 1-\alpha$. When $\alpha = 1$, we recover the original Shapley procedure. In the focal case of $\alpha = 1/2$, the gains are split evenly. For all $0 \leq \alpha < 1$, the existing members benefit when an additional player joins. We think this more general procedure is of interest because, at an intuitive level, it seems fairer, and thus the players may be more willing to go along with such a procedure.

For any $\alpha < 1$, we need to describe how the $1-\alpha$ share is divided up among members of the existing coalition. An obvious approach would be to divide this share equally. However, such a division would not reflect the differing bargaining positions of the members of the coalition. Even a dummy player, one that adds no value, would get the same amount as all the other members. Our approach is to divide the $1-\alpha$ share in stages. Each member has a claim equal to their marginal contribution, where the marginal contribution is to the expanded set consisting of the existing members and the joining player. We begin with equal division, but only up to the smallest claim among the members sharing the gains. After that point, the member with the smallest claim no longer shares in the division. We divide the remaining share equally among the remaining members, but only up to the second-smallest claim. And so on. Finally, if no member of the existing coalition has a claim to the full $1-\alpha$ share, the residual reverts to the joining player. (This case arises only when the joining player has the largest marginal contribution.)

Surprisingly, our procedure yields the Shapley value for all $0 \leq \alpha \leq 1$. We see our result as providing additional support for and understanding of the Shapley value. It shows that the Shapley value arises under more general circumstances, and, arguably, more intuitively appealing ones. It also highlights what is essential to obtaining the Shapley value. The key ingredient is the random ordering of how players join existing coalitions, not the allocation of all the marginal contribution to the joining player.

Dividing up the $1-\alpha$ share among the members of an existing coalition suggests that each member's payoff will depend on the marginal contributions of all the players in the coalition. But the Shapley value payoff to a player depends only on that player's marginal contribution. (This fact characterizes the Shapley value; see Young (1985).) We still get the Shapley value because these extra dependencies cancel out. Each player obtains less than their marginal contribution when joining a coalition. But this shortfall is exactly offset by the extra amounts each player collects when other players join all the coalitions to which that player belongs.

We did not choose our procedure so that this cancellation would occur. We see our procedure as fully justified on intuitive grounds. The fact that it leads back to the Shapley value is a "discovery" and not a feature we built in. Our procedure allows for what we think is a natural interplay between the forces of equal division and bargaining power. The equal division aspect resembles the "principle of the divided cloth" (O'Neill, 1982; Aumann and Maschler, 1985), although we obtain the Shapley value not the nucleolus (Schmeidler, 1969). See Section VI.A for further discussion.

A procedure is a way of understanding a cooperative solution concept, but it is not a non-cooperative implementation of the concept. We do not see this as a shortcoming of our analysis. A procedure is playable in an intuitive sense, where the players agree to various divisions of value and do so without an exact offer-counteroffer protocol that a non-cooperative analysis would require. Nevertheless, it is possible to provide a fully non-cooperative game that mimics our procedure and builds on the non-cooperative implementation of the Shapley value in Hart and Mas-Colell (1996). As we will see, our procedure also suggests a new family of generalized Shapley values, related to but different from the well-known weighted Shapley value (Shapley, 1953).

There are several other computational formulas that lead to the Shapley value. Examples include Harsanyi (1963), Kleinberg and Weiss (1985), Rothblum (1985, 1988), Felsenthal and Machover (1996, 1997), and Bernardi and Freixas (2018). These formulas do not directly translate into a sequential procedure for the division of value. For example, Kleinberg and Weiss (1985) and Rothblum (1985) start by giving each player an equal share of the total and then make adjustments based on how much value coalitions containing a given player create relative to the average coalition. Felsenthal and Machover (1996, 1997) and Bernardi and Freixas (2018) offer a sequential procedure, one in which players join in a random order with each voting randomly yes or no, thus creating $2^n \times n!$ possible orderings. The "yes" voters join the prior yes voters and are given their marginal contribution to that group. The "no" voters are given their marginal contribution to the set absent the prior no voters. It turns out that this procedure is equivalent to the original Shapley procedure, just with a different ordering.[1]

The Shapley procedure and ours build up coalitions by adding an outside player to each existing coalition (in a random order). We also consider a reverse procedure in which we start with the full set of players and then potentially exclude players one at a time (in a random order). For each random construction of a coalition, there is an equivalent deconstruction: instead of players joining in order ABC, say, we can think of first C, then B, then A contemplating being excluded. The reverse procedure can be understood as a scenario in which all players are present at the start and consider what might happen if they fail to reach an agreement. The players would agree to accept the expected value of the exclusion procedure. That expectation is again the Shapley value. The reverse procedure is closest to the non-cooperative implementation of the Shapley value, which we review later.

We begin in Section I by examining how our procedure works in three-person

[1] The procedure is equivalent to the Shapley procedure with a reordering of the joining order in which all the "yes" players join first and then the "no" players join in reverse order. For example, if there are five players who join in order 1(Y), 2(Y), 3(N), 4(Y), 5(N), the payoff to each is identical to an ordering of 1, 2, 4, 5, 3 in the traditional Shapley procedure.

games and show that it leads to the Shapley value. A specific numerical example helps motivate the way that our procedure divides the $1-\alpha$ share of value among the members of a coalition joined. Section II presents our general framework for any number of players and proves that our procedure yields the Shapley value. We follow the proof with an explanation of why our procedure yields payoffs that a priori depend on α, but in the end do not. Section III discusses the question of uniqueness. Section IV modifies our procedure to run in reverse. The reverse procedure also points to a new form of the generalized Shapley value where the chance a player is excluded depends on that player's marginal contribution. We examine our generalized Shapley value in Section V. This generalization is similar to the weighted Shapley value (Shapley, 1953; Kalai and Samet, 1987), but differs in that the weights are endogenous to the game. Section VI contains further discussion of related literature, a non-cooperative implementation of our procedure, and a brief conclusion.

I. Three-Player Examples

We illustrate our procedure (and how it leads to the Shapley value) in the case of three players A, B, and C. Let v denote the characteristic function and set $v(A) = v(B) = v(C) = 0$ (a normalization). For notational simplicity, we write $v(A)$ rather than $v(\{A\})$, and similarly with all sets of players. For $i = A, B, C$, we denote the marginal contribution of player i to coalition S by $m_i(S) = v(S) - v(S \backslash i)$. We order the players by their marginal contributions to the grand coalition.[2] We typically write m_A for $m_A(ABC)$, and we likewise use the abbreviations m_B and m_C. Our ordering implies $m_A \leq m_B \leq m_C$. (In the case of a tie, the ordering can be chosen at random.)

As in the Shapley procedure, players join existing coalitions in a random order. Different from the Shapley procedure, the marginal contribution created by the joining player is divided up in the proportion $\alpha\colon 1-\alpha$, with share α going to the joining member and share $1-\alpha$ going to members of the existing coalition. When $\alpha = 1$, the two procedures coincide.

When $\alpha < 1$, we have to specify how the $1-\alpha$ share is allocated among the members of the existing coalition. As explained in the Introduction, equal division would fail to reflect the players' different bargaining positions. In particular, equal division would mean that a player with zero marginal contribution to any coalition—that is, a dummy player—would get a positive payoff. Instead, our procedure allocates the $1-\alpha$ share equally in stages, with each player inside the coalition being capped by a claim to their marginal contribution. Any unclaimed value left over after this process reverts to the player joining the coalition.

For example, when $\alpha = 0$ and player C joins the coalition $\{AB\}$, the amount to be divided up is m_C. Think of m_C as being split into three parts: $m_C =$

[2] It is a special feature of the three-player case that only the ordering with respect to the grand coalition matters. In the general treatment in Section II, the ordering of the players by marginal contributions to a coalition S will differ across S.

$m_A + (m_B - m_A) + (m_C - m_B)$. Players A and B have an equal claim on the amount m_A, since it does not exceed either of their marginal contributions. Thus m_A is split equally between A and B. Player A's claim stops at m_A, and so player B is the only one with a claim on $m_B - m_A$ and receives this full amount. Neither player A nor player B has a claim on $m_C - m_B$, since player B's claim stops at m_B, and so this value reverts to player C. Notice that this reversion arises only when the player joining the coalition is the one with the largest marginal contribution.

Table 1 shows the payoffs to player A under our procedure, where the row labelled ABC refers to the case that player A arrives first, player B second, and player C third; and likewise for the other rows. We depict the payoffs to player A for the two extreme cases $\alpha = 0$ and 1. The payoffs for $0 < \alpha < 1$ are the corresponding weighted average of the two extreme payoffs.

The average payoff to player A is the same for both extreme cases, so this common value must also be the average value for any $0 < \alpha < 1$. Moreover, the average payoff when $\alpha = 1$ is the Shapley value, by definition. Thus, the average payoff to player A is equal to its Shapley value under our generalized procedure for any value of α.

TABLE 1—PAYOFFS TO PLAYER A

	Payoff to player A when $\alpha = 1$	Payoff to player A when $\alpha = 0$
ABC	0	$m_A(AB) + m_A(ABC)/2$
ACB	0	$m_A(AC) + m_A(ABC)/2$
BAC	$m_A(AB)$	$m_A(ABC)/2$
CAB	$m_A(AC)$	$m_A(ABC)/2$
CBA	$m_A(ABC)$	0
BCA	$m_A(ABC)$	0
Average	$[m_A(AB) + m_A(AC) + 2m_A(ABC)]/6$	$[m_A(AB) + m_A(AC) + 2m_A(ABC)]/6$

An analogous calculation for player B establishes that this player, too, receives its Shapley value under our procedure for any α. Turning to player C, who has the highest marginal contribution, we have to take into account the potential reversion of value to this player when the claims of players A and B are exhausted. This makes the calculation for player C more complicated. But we can use the fact that the total value $v(ABC)$ is allocated under both the Shapley procedure and our procedure. If players A and B receive their Shapley value then player C, too, must also get its Shapley value.

The essential feature of our procedure is the nature of the division of the $1 - \alpha$ share among the members of the existing coalition. There are two steps to how this division works. The first step is that players' claims are limited by their marginal contributions. We assume this not just to ensure that dummy players get zero. The justification for this condition goes back to the core (Gillies, 1959) and is standard.

Our procedure adds a second step, which stipulates equal division to the players up to their marginal contributions. For example, suppose there is a total of 2 to go around and players A and B have claims of 1 and 2, respectively. Then the "first" interval of value, viz. 1, is divided equally since both players have a claim on this quantity. The "second" interval of value, which is again 1, goes to player B, since only the second player has a claim on this interval. More generally, equal division applies to successive intervals of value, where players drop out once their marginal-contribution bounds are exceeded.

To help develop further intuition for this equal-division step, consider an example from Littlechild and Owen (1973), where multiple airlines share the cost of a common runway. Suppose airline A requires a runway of length 1, while airlines B and C each require a runway of length 2. The cost per unit length is 1. The airlines save a total of 3 by sharing a common runway. The characteristic function v for this game is:

$$v(A) = v(B) = v(C) = 0; \quad v(AB) = v(AC) = 1, \quad v(BC) = 2; \quad v(ABC) = 3.$$

The marginal contribution of airline A is $m_A = 1$, while the marginal contributions of airlines B and C are $m_B = m_C = 2$. In this example, the marginal contributions are the cost savings, and it is more intuitive to refer to them as such.

Suppose $\alpha = 1/2$ and the arrival order is ABC. Equal division, constrained by marginal contributions, is natural in this context. Initially, airline A pays for the full cost of 1 for the first leg of the runway. When airline B joins, it pays for half the cost of the first leg (since $\alpha = 1/2$) and pays the full cost of 1 for the second leg. The key question is how $\{AB\}$ should then divide their half share of the cost saving of 2 when airline C joins $\{AB\}$. Airline A has only paid for a share of the first leg and only uses the first leg; therefore, it has no claim to cost savings associated with the second leg. Airlines A and B split equally the savings associated with the first leg of the runway. Airline B keeps all the savings associated with the second leg of the runway. Thus, airline A saves $1/2 \times (1/2 \times 1) = 1/4$, and airline B saves $1/2 \times (1/2 \times 1 + 1 \times 1) = 3/4$. The two players being joined share equally in the division of the $1 - \alpha$ portion of value created, where $\alpha = 1/2$ in this example, but only up to their respective marginal contributions. We think this is a fair division of the cost savings between the two airlines.

II. The α-Procedure

We now define our general procedure for transferable-utility (TU) cooperative games and show that the expected payoffs to the players are independent of α.[3] This immediately implies that our general procedure yields the Shapley value, since the two procedures are identical when $\alpha = 1$.

[3] Brandenburger and Nalebuff (2023*a*) extend the α-procedure to NTU games.

Fix a player set N. For each subset $S \subseteq N$, there is a real number $v(S)$ which is the value created by the players in S. We choose the normalization $v(i) = 0$ for all i, and we assume that v is super-additive; that is, if $S \cap T = \emptyset$, then $v(S \cup T) \geq v(S) + v(T)$.

Let the coalition S without player $i \in S$ be denoted by $S\backslash i$. Denote the marginal contribution of a player i in S to coalition S by $m_i(S) = v(S) - v(S\backslash i)$. By super-additivity, $m_i(S) \geq 0$. Let the number of players in coalition S be denoted by $|S|$. We adopt the following labeling convention: for each coalition S, order the players according to their marginal contributions to S, so that $m_1(S) \leq m_2(S) \leq \cdots \leq m_{|S|}(S)$. If two players have equal marginal contributions, then order them arbitrarily. Note that this ordering will generally be different across different coalitions S. To reduce notion, we often refer to $m_i(S)$ by m_i when the meaning is clear.

In our procedure, like the Shapley procedure, players join existing coalitions in a random order. Unlike the Shapley procedure, the gains are split $\alpha\colon 1-\alpha$ between the joiner and the existing coalition members, where $0 \leq \alpha \leq 1$. Instead of summing up across all $|N|!$ possible orderings, we focus on the $|S|$ possible orderings for the last step in how set S can be formed.

In the Shapley procedure, each player $i \in S$ receives $m_i(S)$ when joining $S\backslash i$ and receives zero in the other $|S-1|$ cases where a player $k \in S\backslash i$, joins i in $S\backslash k$. On average, player i receives $m_i(S)/|S|$. In our procedure, each player $i \in S$ receives $\alpha m_i(S)$ when joining $S\backslash i$ and receives a share of $(1-\alpha)m_k(S)$ when player k joins $S\backslash k$.

It remains to specify the division of the amount $(1-\alpha)m_i$ among the $|S|-1$ players in $S\backslash i$. We divide the $(1-\alpha)m_i$ value equally among the players who have a claim to it. To be precise, player i's marginal contribution m_i is divided into i intervals $(0, m_1]$, $(m_1, m_2]$, $\ldots$, $(m_{i-1}, m_i]$. The kth interval $m_k - m_{k-1}$ is then divided equally among the players with $m_j \geq m_k$. There are $|S|-k$ players who have a claim to that interval. Thus, player $j \in S\backslash i$ gets its share on each interval up to the smaller of its maximum claim m_j and the total amount m_i to go around. Formally, player j receives:

$$(1-\alpha) \sum_{k=1}^{\min(i,j)} \frac{m_k - m_{k-1}}{|S|-k}. \tag{1}$$

In the first term of the sum, we introduced a fictitious player 0 who is not a member of any coalition, and we set $m_0(S) = 0$ for all S. This is done to simplify the notation.

Start with player 1, who is the player with the lowest marginal contribution. This player gets $(1-\alpha)m_1/(|S|-1|)$ no matter who the outside player is. This exhausts player 1's claim, and the remaining value is split among one fewer players. Player 2 gets the same $(1-\alpha)m_1/(|S|-1)$ as player 1 plus an additional $(1-\alpha)(m_2 - m_1)/(|S|-2)$ when the joiner is anyone other than player 1. Here, the amount $m_2 - m_1$ is divided equally among the $|S| - 2$ players whose marginal contributions exceed m_1. This procedure continues for all subsequent players in $S \backslash i$.

When the joining player is $i = |S|$, the procedure works just as before, except that none of the players $j \in S \backslash |S|$ have a claim on the final interval of value $m_{|S|} - m_{|S|-1}$. This amount therefore reverts to player $|S|$. We can see this last step in equation (1), since for any player $j \in S \backslash |S|$, we have $\min(|S|, j) = j < |S|$ and so the interval $m_{|S|} - m_{|S|-1}$ does not appear in the sum for any player j.

The players in $S \backslash i$ have "priority" over the amount $(1-\alpha)m_i$. These players receive this full amount except when $i = |S|$, in which case there is a final interval of value that reverts to i. Putting everything together, our procedure specifies that when player i joins the coalition $S \backslash i$, player j gets a payoff $\pi_{j|i}$ given by:

$$\pi_{j|i}(S) = \begin{cases} \alpha m_j & \text{for } i = j < |S|; \\ \alpha m_{|S|} + (1-\alpha)(m_{|S|} - m_{|S|-1}) & \text{for } i = j = |S|; \\ (1-\alpha) \displaystyle\sum_{k=1}^{\min(i,j)} \frac{m_k - m_{k-1}}{|S| - k} & \text{for } i \neq j. \end{cases} \tag{2}$$

We call this the **α-procedure**. A feature of this procedure is that no player gets more than its marginal contribution to a coalition. In particular, a dummy player gets zero.

THEOREM 1: *In the α-procedure, the expected payoffs to all players are independent of α in $[0, 1]$ and are equal to the Shapley values.*

PROOF: Fix a coalition S, and consider the expected payoff $\pi_j(S)$ to player $j \in S$. This amount is the average over the amount j receives when joining $S \backslash j$ and the amounts j receives across all $|S| - 1$ cases where a player $i \in S \backslash j$ joins player j in an existing coalition. We can write:

$$\pi_j(S) = \frac{1}{|S|} \sum_{i \in S} \pi_{j|i} = \frac{1}{|S|} \sum_{i \in S} \pi_{i|j} = \frac{m_j(S)}{|S|}. \tag{3}$$

The second equality holds because $\pi_{j|i} = \pi_{i|j}$, which follows from the fact that the indices i and j enter entirely symmetrically into equation (2). The third equality holds because, in the α-procedure, player j's marginal contribution is fully allocated across the players $i \in S$.

Equation (3) says that for every coalition S, the expected payoff to any player is independent of α. The same is therefore true when summing up over coalitions. But the α-procedure coincides with the Shapley procedure when $\alpha = 1$. Thus, it must also coincide with the Shapley procedure for all α. ■

We find this result quite surprising. A player's Shapley value is a function only of that player's marginal contributions in the game. In the α-procedure, when player i joins the coalition $S \backslash i$, the payoff to player j depends, prima facie, on a set of marginal contributions, as in equation (2). However, in the expected value calculation for j, the dependence on these other marginal contributions goes away, leaving only the dependence on j's marginal contribution.

The key observation is that what a player gives up when joining a coalition is exactly offset by what that player gets when it is inside the coalition and each of the other players joins. Take player 1. It gives up $(1-\alpha)m_1$ when it joins, but it gets back $(1-\alpha)m_1/(|S|-1)$ a total of $|S|-1$ times—once for each of the other $|S|-1$ players that joins it in S. The loss and gains exactly offset each other. Player 2 gives up $(1-\alpha)m_2$ when it joins. It gets $(1-\alpha)m_1/(|S|-1)$ when player 1 joins it in S, and it gets $(1-\alpha)m_1/(|S|-1)+(1-\alpha)(m_2-m_1)/(|S|-2)$ every time one of players 3 through $|S|$ joins it in S. Again, what player 2 gives up when joining it gets back when others join it. The sacrifices and gains exactly cancel out for players 1, 2, and all the others.

We did not define the α-procedure with the goal of arriving at the Shapley value. Instead, we looked at the problem of how to allocate a coalition's payoff among its members in a defensible manner. This makes the α-procedure more complicated than the Shapley procedure. But it is a natural way to treat members of a coalition equally, subject to respecting the marginal-contribution bound of each player. And when the joiner is player $|S|$, so that no one has a claim on the last interval, the procedure re-allocates this value back to player $|S|$.

Theorem 1 offers new insight into the Shapley value, which can now be seen as arising with a more balanced split between the joiner and the coalition of players being joined. The key assumption is not that the joiner gets everything, as in the original Shapley procedure, but that the order in which players join coalitions is random.

III. Uniqueness

Having established that the α-procedure leads to the Shapley value, it is natural to ask if other procedures also do so.[4] The answer is a qualified yes and we provide one such example below. These other procedures, however, have properties that may be considered unreasonable.

Consider a more general division of the $1-\alpha$ share of value among the members of a coalition. As before, when player i joins $S\backslash i$, we divide i's marginal contribution m_i into i intervals $(0,m_1]$, $(m_1,m_2]$, $\ldots$, $(m_{i-1},m_i]$, all multiplied by $1-\alpha$. For each $k\leq i$, we specify weights $w_{jk}(S\backslash i)\geq 0$ that determine each player j's share of of the kth interval of value $(m_{k-1},m_k]$.

In our α-procedure, the weights, which we write as $w^{\alpha}_{jk}(S\backslash i)$, are given by

$$(4)\qquad w^{\alpha}_{jk}(S\backslash i)=\begin{cases}\dfrac{1}{|S|-k} & \text{for } j\neq i, k\leq \min(i,j);\\ 1 & \text{for } i=j=k=|S|;\\ 0 & \text{otherwise.}\end{cases}$$

[4]We thank Sergiu Hart for suggesting we consider this question.

These weights are constant across i and j among the players with a claim to that interval. In general, the weights could vary across i and j and not be restricted to players $j \geq k$. We now provide a three-player example with unequal weights that leads to the Shapley value. Weights for the first interval $w_{j1}(123\backslash i)$ are

$$
\begin{aligned}
w_{11}(123\backslash 2) &= 1, \qquad & w_{11}(123\backslash 3) &= 0, \\
w_{21}(123\backslash 1) &= 0, \qquad & w_{21}(123\backslash 3) &= 1, \\
w_{31}(123\backslash 1) &= 1, \qquad & w_{31}(123\backslash 3) &= 0.
\end{aligned}
$$

All other weights are given by the w^α in Equation (4).

Focus on the first interval. When player 2 joins $\{13\}$, player 1 gets $(1-\alpha)m_1$ and player 3 gets 0. When player 3 joins $\{12\}$, player 1 gets 0 and player 2 gets $(1-\alpha)m_1$. When player 1 joins $\{23\}$, player 2 gets 0 and player 3 gets $(1-\alpha)m_1$.

Across these three cases, each player gets $(1-\alpha)m_1$, the same as under the w^α-weights. All other payoffs match the w^α results since the weights are the same. Hence, the result is still the Shapley value. What is counterintuitive about this example is the relationship between bargaining strength and allocations. Recall that players are ordered by their marginal contributions, with player 1 having the lowest contribution. And yet, when player 2 joins $\{13\}$, the weaker player 1 ends up with the full amount of m_1 while the stronger player 3 gets none of m_1.

This counterintuitive aspect of this example is why we wrote a qualified yes above. In a companion paper we show that the α-procedure uniquely leads to the Shapley value among all weighing systems that satisfy Monotonicity and Priority; see Brandenburger and Nalebuff (2023*b*). Monotonicity rules out the example above. Stronger players should get at least as much as weaker players. Specifically, players with greater marginal contributions should have weights on each interval that are at least as big as players with lower marginal contributions.

The other axiom we employ is Priority, which gives the players $j \in S\backslash i$ precedence in dividing up the $(1-\alpha)m_i$ share. The very nature of the $\alpha\colon 1-\alpha$ division is that the players in the coalition being joined should be the ones dividing up the $(1-\alpha)m_i$. The only case in which the players being joined do not receive this full amount is when the largest marginal contribution among players in $S\backslash i$ falls short of m_i. This arises when $i = |S|$, and then no player $j \in S\backslash i$ has a claim on the last interval of value $m_{|S|} - m_{|S|-1}$, so this last interval reverts to player $|S|$.

IV. A Reverse Procedure

Our α-procedure is similar to the Shapley procedure in that value is divided step-by-step as players build up to the value $v(N)$ of the grand coalition. Here, we examine what happens if, instead, the players start with the grand coalition and then consider what would happen if they fail to reach an agreement on dividing $v(N)$. In this "reverse" procedure, the players analyze potential break-up scenarios and employ backward reasoning to calculate expected payoffs and the implied division of $v(N)$.

Apart from intrinsic plausibility, there are two additional reasons for looking at a reverse procedure. The reverse procedure naturally suggests a non-cooperative implementation of the Shapley value, analogous to the one in Hart and Mas-Colell (1996); see Section VI.B. Also, our reverse procedure suggests a new generalization of the Shapley value, discussed in Section V.

In the reverse procedure, we think of the players as negotiating in the shadow of an elimination process that operates if they do not reach an agreement regarding the division of $v(N)$. Specifically, if agreement is not reached, some player $i \in N$ is chosen at random to be at risk of being eliminated from the coalition and thereby getting zero in subsequent divisions. This player i negotiates with the other players over the payoff it will accept in order to remain in the coalition. The reverse procedure determines this payoff by assigning player i an α share of its marginal contribution to $v(N)$, while the other players split the $1-\alpha$ share of this marginal contribution in exactly the same fashion as our earlier α-procedure. The expected payoff $\pi^j(N)$ to player j is then given by:

$$\pi^j(N) = \frac{1}{|N|} \sum_{i \in S} \left[\pi_{j|i}(N) + \pi^j(N \backslash i) \right], \tag{5}$$

where $\pi_{j|i}(N)$ specifies the $\alpha\colon 1-\alpha$ split just as in Equation (2) (applied to $S = N$) and $\pi^j(N \backslash i)$ is the anticipated expected payoff to player j in the game without i.

To determine $\pi^j(N \setminus i)$, we repeat the hypothetical elimination process. A player $h \in S \backslash i$ is chosen at random to be at risk of elimination. There is then an $\alpha\colon 1-\alpha$ division of h's marginal contribution just as before. This hypothetical elimination process moves to the set of player $S \backslash \{i, h\}$ and continues in analogous fashion until only one player remains. The resulting backward-recursion equation that determines the allocation of value is based on the same $|N|!$ potential payoffs in our original α-procedure. Since the expected payoffs are identical whether run forwards or backwards, the reverse procedure generates the Shapley value.

V. Generalized Shapley Values

Shapley (1953) and Kalai and Samet (1987) extended the original Shapley value to incorporate weights that depend on the identity of a player. These weights can be understood as influencing the order in which players join existing coalitions. For example, suppose player A has weight 2 and player B has weight 1. Then, the value $v(AB)$ will be divided 2: 1 in favor of player A over player B. In effect, the joining order BA is twice as likely as the order AB, so that player A gets $2/3 \times v(AB)$ in expected value, while player B gets only $1/3 \times v(AB)$ in expected value.

Our generalized procedure suggests a different way to generalize the Shapley value via weights. In the reverse procedure, we supposed that all players face an equal chance of elimination at each stage. More generally, some players might face a higher probability of elimination than others. For example, if a coalition breaks down, it might do so in the least expensive way, by which we mean that the player(s) with the lowest marginal contribution would be the first to be at risk of elimination.

Return to our runway example of Section I. In a weighted variant of that game, if the coalition $\{ABC\}$ breaks down, airline A would be the first to be at risk of elimination. This is because less value is lost if airline A is eliminated as compared with excluding airlines B or C. This line of argument leads to two equally likely elimination orderings: A then B, or A then C. The resulting expected payoffs to the airlines A, B, and C, are α, $(1-\alpha)/2+1$, and $(1-\alpha)/2+1$, respectively.

To see this, note that airline A would be at risk first and therefore earns a payoff of $\alpha \times 1$, namely, the share of marginal contribution that goes to the at-risk party. Airlines B and C each receive $(1-\alpha)/2$ from this stage, and they receive another 2α or $2(1-\alpha)$ depending on which of them is next at risk. Since each order of potential elimination is equally likely, the expected value from this second stage is 1. In the spirit of our reverse procedure, the three airlines would agree to this division at the outset.

Unlike the case where all orders of elimination are equally likely, the expected payoffs now depend on the parameter α. Airline A might do better or worse than its Shapley value payoff of 2/3. This makes sense. If $\alpha = 1$, it is advantageous to be the first airline subject to elimination. If $\alpha = 1/2$, so that the at-risk airline's marginal contribution is split equally, airline A gets 1/2, which is worse than under the Shapley value.

We believe that our notion of a weighted Shapley value establishes a reasonable range of allocations, which is determined by considering a range of reasonable weights. If we fix $\alpha = 1/2$ as a focal case, then, in the runway example, airline A's payoff lies in the interval $[1/2, 2/3]$. The best case for A arises under random order of elimination, in which case A pays 1/3 of the first leg's cost and therefore nets 2/3. The worst case for A arises when it would be the first to be eliminated. In this case, airlines B and C have in effect pre-formed an alliance. Airline A is negotiating with $\{BC\}$, and therefore pays 1/2 and nets 1/2.

The weighted procedure can also be run in a forward direction, like our original α-procedure. Now, the first pair of airlines to form would be the one that creates the largest value, namely, airlines B and C. Airline A joins last. The fact that the ordering is the same in either direction is a consequence of there being only three players in the game. The forward and reverse weighted procedures will generally differ once there are four or more players. It is possible that the pair of players that creates the largest value—and thus would join first in the forward procedure—could include a player with the lowest marginal contribution to the grand coalition, and would therefore be the first to be at risk in the reverse procedure. In general, the payoffs depend both on the endogenous weights and on whether the procedure runs forwards or in reverse.

Summing up, the weights we introduce depend on the players' marginal contributions and are therefore endogenous to the game. Note two features. First, our weights are anonymous. Thus, power arises from a player's contributions in a game, not from their identity. Second, we preserve an equal split in any two-player negotiation. This follows because two players always have the same marginal contribution in any two-person game. We think symmetry in the two-player case is natural. Power in a negotiation should come from some asymmetry—which cannot arise, at least endogenously, in a two-party negotiation. As we vary the weights chosen, our generalized scheme predicts a range of potential outcomes, where the natural boundary cases are a deterministic ordering (in order of increasing marginal contributions) and a completely random ordering. Outcomes outside this range may be seen as less reasonable.

VI. Discussion

A. Related Work

We noted the connection with O'Neill (1982) and Aumann and Maschler (1985). In the division between the first two players when a third player joins, we follow the principle of the divided cloth. In general, though, our analyses differ. Aumann and Maschler obtain the nucleolus, while we obtain the Shapley value. In Aumann and Maschler, the key step in extending the principle of the divided cloth to any number of players is to impose a consistency axiom relating larger games to smaller games (as is commonly done in cooperative game theory). Our α-procedure can be thought of as an alternative extension of the principle of the divided cloth to any number of players—an extension that imposes equal division constrained by marginal contributions.

Owen (1977) proposed a modified game in which some coalitions are more likely to form than others (or come pre-formed). He assumes that when a coalition $\{AB\}$ negotiates with player C, the gain is split evenly, which is akin to setting $\alpha = 1/2$ in our formulation. The value $v(AB) + m_C(ABC)/2$ is then divided between A and B in a manner that preserves the payoff difference between A and B in the Shapley value. Our approach is different since the Shapley value is an output not an input in our analysis.

Aumann and Drèze (1974, Section 12.6) discuss the order in which coalitions might form in a cooperative game. Some coalitions would be likely to form first because of the large value they create for their participants. This is a similar motivation to the non-random order of coalition formation (or breakup) presented in Section V. Unlike our approach, Aumann and Drèze allow a coalition to break up were another player to join the negotiation.

B. Non-Cooperative Implementation

There is a longstanding program to find non-cooperative implementations of cooperative solution concepts. Gul (1989, 1999) initiated this program in the case of the Shapley value, with further progress made by Winter (1994), Hart and

Mas-Colell (1996), Hart and Levy (1999), Serrano (2005, 2021), and McQuillin and Sugden (2016).

We can convert our reverse α-procedure into a non-cooperative game following closely the approach in Hart and Mas-Colell (1996). The α-procedure provides a division of value between the player at risk of exclusion and the other players. The same division can be obtained via a suitably defined non-cooperative game. A player i is selected to be at risk of exclusion. A player j (possibly i) is then selected to be the proposer and make an ultimatum offer to the other players. Player j claims the full marginal contribution of player i and offers everyone else (player j included) their Shapley value in the game without i. In equilibrium, such proposals are always accepted. Each player's chance of being selected as the proposer follows Equation (2) from Section II: the chance j makes the proposal when i is at risk is $\pi_{j|i}(S)/m_i$. In particular, player i has an α chance of being the proposer and the remaining $1-\alpha$ probability is allocated among the other players (except for the reversion case). When all players have an equal chance of being at risk of exclusion, this non-cooperative game is the analog to our reverse α-procedure, and the stationary subgame-perfect equilibrium payoffs are equal to the Shapley values. This game differs from Hart and Mas-Colell in two main ways. First, player j knows which player i is at risk before proposing, and, second, the proposer selection probabilities depend on the marginal contributions.

C. Conclusion

We introduce a new bargaining procedure that generalizes the original Shapley (1953) procedure in sharing gains in an $\alpha : 1-\alpha$ manner between the player joining a coalition and the members of that coalition. Surprisingly, our α-procedure also leads to the Shapley value. Though somewhat more complicated than the Shapley procedure, we think the α-procedure is intuitive. It allows for a more balanced split between joiner and coalition, and it allocates value among members of the coalition equally, subject to respecting the marginal-contribution bound of each player in the coalition.

The α-procedure also suggests a new solution concept that differs from the weighted Shapley value in allowing endogenous weights that depend on marginal contributions. This approach is motivated by the observation that some coalitions might be more likely to form than others, depending on the value they create.

We believe that our α-procedure—especially the focal case of $\alpha = 1/2$ rather than the original Shapley case of $\alpha = 1$—adds an important justification to the use of the Shapley value as a recommended division of value to the players in a game.

REFERENCES

Aumann, Robert J, and Jacques H Drèze. 1974. "Cooperative Games with Coalition Structures." *International Journal of Game Theory*, 3(4): 217–237.

Aumann, Robert J, and Michael Maschler. 1985. "Game Theoretic Analysis of a Bankruptcy Problem from the Talmud." *Journal of Economic Theory*, 36(2): 195–213.

Bernardi, Giulia, and Josep Freixas. 2018. "The Shapley Value Analyzed under the Felsenthal and Machover Bargaining Model." *Public Choice*, 176(3–4): 557–565.

Brandenburger, Adam, and Barry Nalebuff. 2023*a*. "Reevaluating the Shapley Value: The NTU Case." Available at adambrandenburger.com/aux/material/NTU.

Brandenburger, Adam, and Barry Nalebuff. 2023*b*. "Reevaluating the Shapley Value: Uniqueness of the α-Procedure." Available at adambrandenburger.com/aux/material/Unique.

Brock, Horace W. 1992. "Game Theory, Symmetry, and Scientific Discovery." In *Rational Interaction: Essays in Honor of John C. Harsanyi*. 391–418. Berlin: Springer-Verlag.

Felsenthal, Dan S, and Moshé Machover. 1996. "Alternative Forms of the Shapley Value and the Shapley–Shubik Index." *Public Choice*, 87(3–4): 315–318.

Felsenthal, Dan S, and Moshé Machover. 1997. "Ternary Voting Games." *International Journal of Game Theory*, 26: 335–351.

Gillies, Donald B. 1959. "Solutions to General Non-Zero-Sum Games." In *Contributions to the Theory of Games IV*. Vol. 40 of *Annals of Mathematics Studies*, Ed. A. W. Tucker and R. D. Luce, 47–85. Princeton: Princeton University Press.

Gul, Faruk. 1989. "Bargaining Foundations of Shapley Value." *Econometrica*, 57(1): 81–95.

Gul, Faruk. 1999. "Efficiency and Immediate Agreement: A Reply to Hart and Levy." *Econometrica*, 67(4): 913–917.

Harsanyi, John C. 1963. "A Simplified Bargaining Model for the n-Person Cooperative Game." *International Economic Review*, 4(2): 194–220.

Hart, Sergiu, and Andreu Mas-Colell. 1996. "Bargaining and Value." *Econometrica*, 64(2): 357–380.

Hart, Sergiu, and Zohar Levy. 1999. "Efficiency Does Not Imply Immediate Agreement." *Econometrica*, 67(4): 909–912.

Kalai, Ehud, and Dov Samet. 1987. "On Weighted Shapley Values." *International Journal of Game Theory*, 16(3): 205–222.

Kleinberg, Norman L, and Jeffrey H Weiss. 1985. "A New Formula for the Shapley Value." *Economics Letters*, 17(4): 311–315.

Littlechild, Stephen C, and Guillermo Owen. 1973. "A Simple Expression for the Shapley Value in a Special Case." *Management Science*, 20(3): 370–372.

McQuillin, Ben, and Robert Sugden. 2016. "Backward Induction Foundations of the Shapley Value." *Econometrica*, 84(6): 2265–2280.

O'Neill, Barry. 1982. "A Problem of Rights Arbitration from the Talmud." *Mathematical Social Sciences*, 2(4): 345–371.

Owen, Guilliermo. 1977. "Values of Games with a Priori Unions." In *Mathematical Economics and Game Theory: Essays in Honor of Oskar Morgenstern*. Ed. Rudolf Henn and Otto Moeschlinm, 76–88. Berlin: Springer-Verlag.

Rothblum, Uriel G. 1985. "A Simple Proof for the Kleinberg-Weiss Representation of the Shapley value." *Economics Letters*, 19(2): 137–139.

Rothblum, Uriel G. 1988. "Combinatorial Representations of the Shapley Value Based on Average Relative Payoffs." In *The Shapley Value: Essays in Honor of Lloyd S. Shapley*. Ed. A. Roth, 121–126. Cambridge: Cambridge University Press.

Schmeidler, David. 1969. "The Nucleolus of a Characteristic Function Game." *SIAM Journal of Applied Mathematics*, 17: 1163–1170.

Serrano, Roberto. 2005. "Fifty Years of the Nash Program, 1953–2003." *Investigaciones Económicas*, 29: 219–258.

Serrano, Roberto. 2021. "Sixty-Seven Years of the Nash Program: Time for Retirement?" *SERIEs: Journal of the Spanish Economic Association*, 12(1): 35–48.

Shapley, Lloyd S. 1953. "A Value for n-Person Games." In *Contributions to the Theory of Games II*. Ed. H. Kuhn and A. Tucker. Princeton: Princeton University Press.

Winter, Eyal. 1994. "The Demand Commitment Bargaining and Snowballing Cooperation." *Economic Theory*, 4: 255–273.

Young, H Peyton. 1985. "Monotonic Solutions of Cooperative Games." *International Journal of Game Theory*, 14: 65–72.