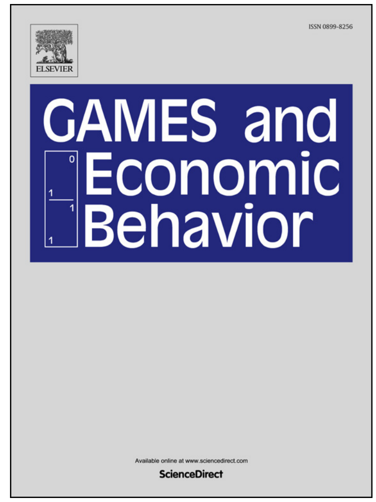
Accepted Manuscript



Balanced externalities and the Shapley value

Ben McQuillin, Robert Sugden

PII: S0899-8256(18)30037-X
DOI: https://doi.org/10.1016/j.geb.2018.03.006
Reference: YGAME 2844

To appear in: *Games and Economic Behavior*

Received date: 22 August 2016

Please cite this article in press as: McQuillin, B., Sugden, R. Balanced externalities and the Shapley value. *Games Econ. Behav.* (2018), https://doi.org/10.1016/j.geb.2018.03.006

Balanced externalities and the Shapley value

Ben McQuillin[a], Robert Sugden[b]

School of Economics and Centre for Behavioural and Experimental Social Science, University of East Anglia, Norwich NR4 7TJ, UK.

[a]*b.mcquillin@uea.ac.uk*
[b]*r.sugden@uea.ac.uk*

Abstract

We characterize the Shapley value using (together with standard conditions of efficiency and equal gains in two-player games) a condition of 'undominated merge-externalities'. Similar to the well-known 'balanced contributions' characterization, our characterization corresponds intuitively to 'threat points' present in bargaining. It derives from the observation that all semivalues satisfy 'balanced merge-externalities'. Our characterization is applicable to useful, narrow sub-classes of games (including monotonic simple games), and also extends naturally to encompass games in partition function form.

Keywords:
Shapley value, balanced contributions, merge-externalities, semivalues, coalitional bargaining
JEL classification: C71, C78

1. Introduction

In this paper, we provide a new characterization of the Shapley value (Shapley, 1953). The distinctive axiom of this characterization can be interpreted as a condition of balance between the threats that players can make to one another in an unmodelled bargaining process.

A motivation for our characterization lies in our interpretation of Shapley's own understanding of his contribution. Shapley (1953) presents the value as a proposal about how players will 'evaluate ... the prospect of having to play a game', and treats this as an answer to a foundational problem in the theory of games. The implication is that Shapley's project is to find general principles that characterize the expected utility outcomes of rational play in abstract games. Shapley and Shubik (1969) later describe the value as addressing 'the idea of "fair division" in a socio-economic situation', but they continue by explaining: 'This solution seeks to evaluate each person's position in the game a priori, taking into account both his own strategic opportunities *and his bargaining position*' (our italics). In other words, Shapley and Shubik see the value as also a normative proposition, based on a conception of 'fairness' that seems to correspond with the expected outcome of bargaining between rational individuals – a conception that is characteristic of social contract theory.[1] Taking this perspective, it is natural to look for bargaining foundations for the Shapley value. One approach is to model bargaining as a noncooperative game. Various authors have taken this approach, showing for specific models of the bargaining process that – at least under some conditions – the Shapley value matches utility expectations (Gul, 1989; Hart and Mas-Colell, 1996; Pérez-Castrillo and Wettstein, 2001; McQuillin and Sugden, 2016). In this paper we adopt a different strategy, by aiming to explicitly capture the notion of a player's bargaining position within an axiomatic characterization.

Of the many existing characterizations of the Shapley value, some already can be interpreted as offering a defence of the claim that the value would be the outcome of rational bargaining. In particular, we would suggest that the 'balanced contributions' characterization due to Myerson (1980) does this. Our characterization, based on an idea of 'merge-externalities', shares with the balanced contributions characterization the fact that it references subgames that could be thought of as 'threat points' and could plausibly be used as such by players themselves bargaining over a prospective coalitional surplus. The balanced contributions characterization relates closely to subsequent noncooperative models (Hart and Mas-Colell, 1996; Pérez-Castrillo and Wettstein, 2001) in which there is a possibility, during bargaining, that players leave the game. Our 'merge externalities' characterization relates similarly to models (Gul, 1989; McQuillin and Sugden, 2016) in which, during bargaining, players have opportunities to form intermediate coalitions instead of forming the grand coalition in a single step.

Our characterization also shares with the balanced contributions characterization a useful virtue of applicability to narrow sub-classes of cooperative games. Our characterization *complements* the balanced contributions characterization by the fact that it is applicable to *different* sub-classes of games. In particular, our conditions can also be used to characterize (on the sub-

[1] In social contract theories such as those of Rawls (1971) and Binmore (1998), social states are defined to be just if and only if they would be the outcome of rational bargaining in a hypothetical 'original position'. Principles of fairness enter such theories through the specification of the original position, but the contracting parties are assumed to make rational use of whatever bargaining opportunities they are allowed.

class of simple monotonic games) the measure of voting power proposed by Shapley and Shubik (1954), now known as the *Shapley-Shubik power index*. In addition, our characterization has an advantage of being readily extensible to games in partition function form (first proposed by Thrall and Lucas, 1963), which are applicable to modelling, for example, collusion and merger in imperfectly competitive industries, research alliances, trade blocs, and international environmental agreements.

To fix ideas, consider the 'glove game' (of Shapley and Shubik, 1969) comprising a player set $N = \{1, 2, 3\}$ and a characteristic function $v\colon 2^N \to \mathbb{R}$ such that $v(S) = 1$ wherever coalition S contains player 1 plus at least one other player, and $v(S) = 0$ otherwise. (The intuition is that player 1 starts with a right-handed glove, players 2 and 3 both start with left-handed gloves, and a pair of gloves has a worth of 1.) Suppose players' 'expectations' in this game are given by the imputation (ϕ_1, ϕ_2, ϕ_3).

The balanced contributions idea, applied to this game, runs as follows. Consider three subgames, prospectively formed by one or other player *leaving and taking her glove with her*. The subgames have player sets $N_1 = \{2, 3\}$, $N_2 = \{1, 3\}$, and $N_3 = \{1, 2\}$, and 'expectations', respectively, of $(0, 0)$, $(0.5, 0.5)$ and $(0.5, 0.5)$.[2] For any distinct $i, j, k \in \{1, 2, 3\}$, the consequence for player i of j leaving the game can be measured by the resulting change in i's expectation – that is, i's expectation in the subgame with player set $\{i, k\}$, minus ϕ_i. If players use these subgames as threat points, then the threats (which we could also term 'leave-externalities') become balanced whenever, for each pair of players i, j, the consequence for i of j leaving the game equals the consequence for j of i leaving the game. The balance conditions solve to give $\phi_1 = \phi_2 + \frac{1}{2} = \phi_3 + \frac{1}{2}$. Combining this result with the efficiency property $\phi_1 + \phi_2 + \phi_3 = 1$ gives the Shapley value imputation, $\phi_1 = \frac{2}{3}$, $\phi_2 = \phi_3 = \frac{1}{6}$.

Our 'merge-externalities' approach runs in the following (similar) way. Consider instead the six subgames prospectively formed by some player i *leaving but first giving her glove to some other*. (This event might be interpreted as a 'merger' between i and j that allows j to take full control of the combined resources of the two players.) These subgames have player sets $N_{(1,2)} = \{2^{+1}, 3\}$ (corresponding to 1 having given her resources to 2), $N_{(1,3)} = \{2, 3^{+1}\}$, $N_{(2,1)} = \{1^{+2}, 3\}$, $N_{(2,3)} = \{1, 3^{+2}\}$, $N_{(3,1)} = \{1^{+3}, 2\}$, $N_{(3,2)} = \{1, 2^{+3}\}$, with associated expectations, respectively, of $(1, 0)$, $(0, 1)$, $(1, 0)$, $(0.5, 0.5)$, $(1, 0)$, and $(0.5, 0.5)$.[3] For any distinct $i, j, k \in \{1, 2, 3\}$, the

[2] The expectations for these two-player games can be derived by re-applying the balanced contributions and efficiency principles, or by directly applying a principle that in two-player games players split any coalitional surplus equally.

[3] The expectations for these two-player games are derived by directly applying a prin-

consequence for player i of j leaving the game and giving her resource to k can be measured by the resulting change in i's expectation – that is, i's expectation in the subgame with player set $\{i, k^{+j}\}$, minus ϕ_i. If players use *these* subgames as threat points, then an imbalance between threats exists whenever, for some pair of players i, j, the consequence for i of j leaving the game and giving her resource to the third player k is strictly worse than the consequence for j of i doing the same. The threats are again balanced only by $\phi_1 = \phi_2 + \frac{1}{2} = \phi_3 + \frac{1}{2}$ and hence, with the efficiency property, the Shapley value imputation. We will show that this result generalizes to all transferable utility games in characteristic function form. In games with more than three players one may not want to require that for *every* third party k the consequence for i of j leaving and giving her resource to player k is no worse than that for j of i doing the same ('balanced merge-externalities'), but rather that for *some* third party k this weak inequality holds (we call this 'undominated merge-externalities'). We are able to characterize the Shapley value by combining the condition of 'undominated merge-externalities' with the efficiency condition and the condition that, in two-player games, surplus is divided equally.

It may be noticed that the merge-subgames considered in the argument above have been used previously - originally by Lehrer (1988) and then for example in Nowak (1997) and Casajus (2012) - in characterizations of the *Banzhaf value*.[4] But to the best of our knowledge ours is the first paper to use amalgamation subgames of this type in a characterization of the Shapley value. In characterizing the Banzhaf value, Lehrer (1988) and others refer to the *internal* effect of the amalgamation. Our approach refers instead to the *external* effect.

Normative economics is often based on an underlying conceit of a *social planner*, modelled as an agent – in effect, a benevolent despot – who maximizes her preferred conception of social welfare and is not subject to institutional constraints.[5] Insofar as its recommendations are directed at such an entity, there is arguably no need to give these recommendations any 'bargaining foundation' at all. But in the real world, generally, no such entity exists. The great advantage of the Shapley-Shubik approach is that their

ciple that in two-player games players split any coalitional surplus equally.

[4]Haller (1994) and Malawski (2002), also characterizing the Banzhaf value, use closely-related 'proxy agreement' subgames, differing only in that (after a bilateral merger) an additional null-player also remains.

[5]Buchanan (1987) identifies this conceit when he endorses Wicksell's (1896/1958) argument that 'economists should cease proffering policy advice as if they were employed by a benevolent despot'.

recommendations - so long as they do have a bargaining foundation - can be directed at the bargaining agents themselves. The role of the normative economist need not be so much to 'whisper in the ear' of the social planner, as to advise the bargaining agents that if they were to hold out for a negotiated agreement, the agreement that is now being proposed (in this case, the Shapley value) is the best they could reasonably expect.

In Section 2 below we formalize our merge-externalities characterization of the Shapley value, and show that all 'semivalues' give rise to merge-externalities that are 'balanced'. We also explain why, if 'value' is interpreted in terms of the expectations of rational bargainers, it is natural to impose the condition that surplus is divided equally in two-player games. In Section 3 we show that none of the components of our characterization is redundant. By using a substitute for the efficiency axiom, we can characterize the Banzhaf value; by using alternative substitutes for the two-player axiom, we can characterize the equal division solution and the Shapley value of the homomollifier. (Contrary to the characterizations in other sections, the characterizations in Section 3 are intended only to illustrate formal properties: we do not argue that the combinations of conditions considered in this section are necessarily congruent with any specific conception of bargaining.) In Section 4 we note that in order to interpret our characterization of the Shapley value as a point of balance between *threats*, one might want to require that merge-externalities are negative. By weakening the requirement of undominated merge-externalities to one of undominated *merge-threats* (permitting merge-externalities to be dominated if they are non-negative) we provide a characterization of the Shapley value that holds for specific types of games. It is for these games that the 'bargaining foundation' is most compelling. In Section 5 we discuss the applicability of our characterization to monotonic simple games. In Section 6 we apply the merge-externalities axioms to a solution for games in partition function form.

2. Balanced merge-externalities and the Shapley value

Let Γ denote the set of *transferable utility games in characteristic function form*, generically (N, v) with N denoting a set of players and with $v\colon 2^N \to \mathbb{R}$ having the property $v(\varnothing) = 0$.

We define, for any $(N, v) \in \Gamma$, and for any $i, j \in N$, subgames $\mathcal{L}((N, v), i) \in \Gamma$ and $\mathcal{M}((N, v), (i, j)) \in \Gamma$ using:

$$\begin{aligned} \mathcal{L}((N,v),i) &\equiv (N \setminus \{i\}, v'), \\ \forall S &\subseteq (N \setminus \{i\}), v'(S) = v(S). \end{aligned}$$

$$\begin{aligned}\mathcal{M}((N,v),(i,j)) &\equiv (N\setminus\{i\},v'),\\ \forall S &\subseteq (N\setminus\{i\}),v'(S)=\begin{cases} v(S) & \text{where } j\notin S \\ v(S\cup\{i\}) & \text{where } j\in S\end{cases}.\end{aligned}$$

$\mathcal{L}((N,v),\{i\})$ - which we may term a 'leave-subgame ' - is the subgame that arises naturally from (N,v) if i '*leaves*, taking her resources with her'. In this subgame, relative to (N,v), coalitional payoffs remain unchanged, but coalitions containing player i are no longer possible. $\mathcal{M}((N,v),(i,j))$ - which we may term a 'merge-subgame' - is the subgame that arises naturally from (N,v) if i and j '*merge*' (under the control of j). In this subgame, relative to (N,v), coalitions containing player i are no longer possible but coalitions containing j receive a payoff as if - in the game (N,v) - they also contained i, other coalitional payoffs remaining unchanged. Since it is the *external* effects of the merger in which we are going to be interested, it does not matter to us by what process the 'merge' occurs: the intuition may be that i 'leaves, giving her resources to j' or that i 'is bought-out by j', or that in any other sense i and j combine to form a single, amalgamated entity.[6]

A *solution* (or 'value' in the terminology of Shapley, 1953), generically ϕ, associates with each $(N,v)\in\Gamma$ and each $i\in N$ a real number $\phi_i(N,v)$, in some sense representing i's utility expectation in the game (N,v).

We define some possible conditions on any solution ϕ. The first condition is *efficiency*.[7]

Efficiency. $\forall(N,v)\in\Gamma,\sum_{i\in N}\phi_i(N,v)=v(N)$.

The second condition, *equal gains in 2-player games*, is a weaker variant of a condition often termed (originally by Hart and Mas-Colell, 1989) 'standardness for two-player games'. As noted by Hart and MasColell, this can be straightforwardly derived from more basic conditions: in particular, any solution that (on the domain of 2-player games) satisfies covariance under addition of a constant to one player's utilities (*TU-covariance*) and that assigns symmetric expectations in perfectly symmetric games (*equal treatment of equals*) must satisfy equal gains in 2-player games.

[6]Although, notationally, j 'remains' in the merge subgame, the relevant intuition is just that one player remains instead of two, and that this single entity represents a symmetric amalgam of i and j in the original game. This is the same notion of merging as used in Lehrer (1988).

[7]The efficiency condition is so-named for convenience and by convention, though the name itself is apt only on the sub-class of cohesive games. We have defined Γ to include non-cohesive games, but the results in this paper would all also hold on the sub-class of cohesive games.

TU covariance captures the idea that payoffs players can obtain without cooperation are irrelevant to the division of surplus strictly generated by cooperation. So wherever $(N, \bar{v})$ is an inessential game, $\phi_i(N, \bar{v} + v')$ equals $\bar{v}(\{i\}) + \phi_i(N, v')$. If a 'solution' were interpreted as something to be chosen by a social planner, this condition would express a contestable normative principle. But in the context of bargaining, it expresses a much more intuitive idea – that payoffs that would be achieved whatever the outcome of the bargaining process cannot feature in threats or offers made between ideally rational players.

Any axiomatized solution concept must respect whatever symmetries are built into the theoretical framework that is being used. Thus, if players can be identified only by arbitrary labels (such as $1, 2, \ldots$), equal treatment of equals is unavoidable. In our modelling framework, however, players can sometimes be distinguished by having different histories of previous mergers. (In the glove game example, when the player set is $\{i, k^{+j}\}$, one player has been involved in a previous merger while the other has not.) We implicitly assume that rational behavior at any point in a bargaining process is independent of the history of how that point was reached.

Equal-gains in 2-player games. $\forall (N, v) \in \Gamma, |N| = 2 \rightarrow \forall i, j \in N$,
$\phi_i(N, v) - v(\{i\}) = \phi_j(N, v) - v(\{j\})$.

The next condition, *balanced contributions*, is included here just for purposes of comparative discussion. (For purposes of comparison we could term this 'balanced leave-externalities'.) It specifies a symmetry of consequences: the effect (on utility expectation) for one player i, of another, j, 'leaving', is the same as the effect on j of i leaving.

Balanced contributions (or 'balanced leave-externalities'). $\forall (N, v) \in \Gamma, \forall i, j \in N, \phi_i(\mathcal{L}((N, v), j)) - \phi_i(N, v) = \phi_j(\mathcal{L}((N, v), i)) - \phi_j(N, v)$.

Our remaining conditions also specify symmetries of consequences. In the first of these, *balanced merge-externalities*, for every triple of players, i, j and k, the effect on i of j merging with k is the same as the effect on j of i merging with k.

Balanced merge-externalities. $\forall (N, v) \in \Gamma, \forall i, j, k \in N$,
$\phi_i(\mathcal{M}((N, v), (j, k))) - \phi_i(N, v) = \phi_j(\mathcal{M}((N, v), (i, k))) - \phi_j(N, v)$.

Balanced merge-externalities may be viewed as a rather strong condition. Our next condition, *undominated merge-externalities*, is (prima facie) a much weaker variant of the same idea. It requires only that, for each pair of players

i and j, if there exist further players, there should be *some* third party k such that the effect on i of j merging with k is *no worse than* the effect on j of i merging with k.

Undominated merge-externalities. $\forall(N,v)\in\Gamma, |N|>2\rightarrow\forall i,j\in N$, $\exists k\in(N\setminus\{i,j\}), \phi_i(\mathcal{M}((N,v),(j,k)))-\phi_i(N,v)\geqslant\phi_j(\mathcal{M}((N,v),(i,k)))-\phi_j(N,v)$.

The *Shapley value*, ϕ^{Sh}, is defined using (for any $(N,v)\in\Gamma$, and for any $i\in N$):

$$\phi_i^{Sh}(N,v)=\sum_{S\subseteq N\setminus\{i\}}\frac{|S|!(|N|-|S|-1)!}{|N|!}\left(v(S\cup\{i\})-v(S)\right).$$

The Shapley value belongs to a class of solutions known as *semivalues* (due to Dubey, Neyman and Weber, 1981). A solution ϕ is termed a semivalue when it satisfies, for some probability measure ξ on $[0,1]$, the following:

$$\forall(N,v)\in\Gamma,\forall i\in N,\phi_i(N,v)=\sum_{S\subseteq N\setminus\{i\}}p_{|S|}^{|N|}\left(v(S\cup\{i\})-v(S)\right),$$

where (for any positive integer n and non-negative integer $s<n$)

$$p_s^n\equiv\int_0^1 t^s\,(1-t)^{n-s-1}\,d\xi(t).$$

It is well known (due to Myerson, 1980) that efficiency and balanced contributions together characterize the Shapley value, and it is also known (see Sánchez, 1997) that all other semivalues also satisfy balanced contributions. We report two similar results with respect to balanced merge-externalities.

Proposition 1. *Every semivalue ϕ satisfies balanced merge-externalities.*

Proposition 2. *A solution ϕ satisfies efficiency, equal-gains in 2-player games and undominated merge-externalities if and only if $\phi=\phi^{Sh}$.*

(We prove Propositions 1 and 2 in the Appendix.)

3. Balanced merge-externalities and other solutions

It is straightforward to confirm that there is no redundancy within the characterization expressed in Proposition 2, by briefly noting other well-known solutions. Clearly, the *nucleolus* (Schmeidler, 1969) satisfies efficiency and equal-gains in 2-player games (but not undominated merge-externalities). Also, the *Banzhaf value* satisfies equal-gains in 2-player games and undominated merge-externalities (but not efficiency), and the *equal division value* satisfies efficiency and undominated merge-externalities (but not equal-gains in 2-player games). In this section we formalize these observations as further characterizations.

The Banzhaf value, ϕ^{Bz}, is defined using (for any $(N, v) \in \Gamma$, and for any $i \in N$):

$$\phi_i^{Bz}(N, v) = \sum_{S \subseteq N \backslash \{i\}} \frac{v(S \cup \{i\}) - v\,(S)}{2^{|N|-1}}.$$

Dubey and Shapley (1979) noted that ϕ^{Bz} and ϕ^{Sh} are "fundamentally very similar", in the sense that replacing efficiency, in a conventional characterization of ϕ^{Sh}, with the following condition (here called *Bz-sum*), obtains ϕ^{Bz}.

Bz-sum. $\forall (N, v) \in \Gamma, \sum_{i \in N} \phi_i(N, v) = \sum\limits_{i \in N} \sum\limits_{S \subseteq N} \frac{v(S) - v(S \backslash \{i\})}{2^{|N|-1}}$.

The following proposition reproduces this similarity, using our merge-externalities approach.

Proposition 3. *A solution ϕ satisfies equal-gains in 2-player games, undominated merge-externalities and Bz-sum if and only if $\phi = \phi^{Bz}$.*

(We prove Proposition 3 in the Appendix. Notice that the Banzhaf value is another semivalue, and therefore also satisfies balanced merge-externalities.)

While balanced contributions combined only with efficiency characterizes the Shapley value, our characterizations using merge-externalities have required an additional condition: equal gains in 2-player games. Merge-externalities conditions - unlike balanced contributions - have no bearing in 2-player games, because there are no third-parties to whom merge-externalities can accrue. As previously noted, equal gains in 2-player games derives directly from more primitive conditions: TU-covariance and equal treatment of equals.[8] A simple example of a 2-player game condition that violates TU-covariance (but fulfils equal treatment of equals) has sometimes been termed 'egalitarian standardness for 2-player games'.[9] A weaker variant would be the following condition, *equal-division in 2-player games*.

Equal-division in 2-player games. $\forall (N,v) \in \Gamma, |N| = 2 \rightarrow \forall i,j \in N$, $\phi_i(N,v) = \phi_j(N,v)$.

The equal division solution, ϕ^{ED}, is defined using (for any $(N,v) \in \Gamma$, and for any $i \in N$):

$$\phi_i^{ED}(N,v) = \frac{v(N)}{|N|}.$$

The following result - that efficiency, equal-division in 2-player games, and undominated merge-externalities together characterize the equal division solution - is then very straightforward.[10]

Proposition 4. *A solution ϕ satisfies efficiency, equal-division in 2-player games and undominated merge-externalities if and only if $\phi = \phi^{ED}$. Balanced merge-externalities is also satisfied by $\phi = \phi^{ED}$.*

(We prove Proposition 4 in the Appendix.)

An intuitive example of a 2-player game condition that violates equal treatment of equals (but fulfils TU-covariance) arises if we modify our notion of a merger, to one of *consolidation*: so that if one entity has formed through more mergers than another, then the two need not be treated symmetrically. To represent this idea, for any $(N,v) \in \Gamma$ let $\Pi(N)$ denote the set of all partitions of N. For any $\pi \in \Pi(N)$ we define a *quotient game* $(N,v)^\pi \equiv (\pi, v^\pi)$, with for any $\overline{S} \subseteq \pi$, $v^\pi(\overline{S}) = v\left(\bigcup_{S \in \overline{S}} S\right)$. We fix a game $(N,v) \in \Gamma$ and we then consider the set - which we denote $Q(N,v)$ - of quotient games based on (N,v), $Q(N,v) \equiv \{(N,v)^\pi : \pi \in \Pi(N)\}$. For any $(N,v)^\pi \in Q(N,v)$ and for any $I,J \in \pi$, we define a subgame $\overline{\mathcal{M}}((N,v)^\pi, \{I,J\}) \equiv (N,v)^{(\pi \setminus \{I,J\}) \cup \{I \cup J\}}$, which is the subgame that comes about if I and J consolidate.

---

[8]It is useful also to note that, as it is used in conjunction with either efficiency or Bz-sum, equal gains in 2-player games could be replaced in our characterizations with *balanced contributions in 2-player games*: i.e. $\forall (N,v) \in \Gamma, |N| = 2 \rightarrow \forall i,j \in N, \phi_i(\mathcal{L}((N,v),j)) - \phi_i(N,v) = \phi_j(\mathcal{L}((N,v),i)) - \phi_j(N,v)$.

[9]See for example van den Brink and Funaki (2009).

[10]Notice that while all 'semivalues' satisfy balanced merge-externalities, Proposition 4 demonstrates that the converse is not true. The equal division solution is not a semivalue, but it satsifies balanced merge-externalities nevertheless.

Note that the players of a game in $Q(N,v)$ are represented by sets, and they therefore have a cardinality property (proportionate to the number of mergers through which the player has formed) which is referenced in the following 2-player game condition (defined only on $Q(N,v)$).

Equal per-capita gains in 2-player games. $\forall (\pi, v^\pi) \in Q(N,v), |\pi| = 2 \rightarrow \forall I, J \in \pi, \frac{\phi_I(\pi,v^\pi) - v^\pi(\{I\})}{|I|} = \frac{\phi_J(\pi,v^\pi) - v^\pi(\{J\})}{|J|}$.

To obtain characterizations within this framework, we define conditions (exactly analogous to those related to merge-externalities, but again defined only on $Q(N,v)$) of balanced and undominated consolidation-externalities.

Balanced consolidation-externalities. $\forall (\pi, v^\pi) \in Q(N,v), \forall I, J, K \in \pi$,
$\phi_I\left(\overline{\mathcal{M}}((\pi,v^\pi),\{J,K\})\right) - \phi_I(\pi,v^\pi) = \phi_J\left(\overline{\mathcal{M}}((\pi,v^\pi),\{I,K\})\right) - \phi_J(\pi,v^\pi)$.

Undominated consolidation-externalities. $\forall (\pi, v^\pi) \in Q(N,v), |\pi| > 2 \rightarrow \forall I, J \in \pi, \exists K \in \pi \setminus \{I,J\}, \phi_I\left(\overline{\mathcal{M}}((\pi,v^\pi),\{J,K\})\right) - \phi_I(\pi,v^\pi) \geqslant \phi_J\left(\overline{\mathcal{M}}((\pi,v^\pi),\{I,K\})\right) - \phi_J(\pi,v^\pi)$.

We define $h_v : 2^N \rightarrow \mathbb{R}$ (termed the *homomollifier* of v by Charnes, Rousseau, and Sieford, 1978) using:

$$\forall S \subseteq N, h_v(S) = \frac{|S|}{|N|}(v(N) - v(N \setminus S)) + \frac{|N| - |S|}{|N|} v(S).$$

Efficiency, equal-gains in 2-player games, and undominated consolidation-externalities characterize, on the sub-class $Q(N,v)$, the Shapley value. But efficiency, equal *per-capita* gains in 2-player games, and undominated consolidation-externalities together characterize, on the same sub-class, a solution which can be termed the 'Shapley value of the homomollifier'.[11]

Proposition 5. *(i) A solution ϕ satisfies, on the sub-class $Q(N,v)$, efficiency, equal-gains in 2-player games and undominated consolidation-externalities if and only if for any $(N,v)^\pi \in Q(N,v)$, $\phi((N,v)^\pi) = \phi^{Sh}((N,v)^\pi)$.*

(ii) A solution ϕ satisfies, on the sub-class $Q(N,v)$, efficiency, equal per-capita gains in 2-player games and undominated consolidation-externalities if and only if for any $(N,v)^\pi \in Q(N,v)$, $\phi((N,v)^\pi) = \phi^{Sh}((N,h_v)^\pi)$.

[11] This solution arises, in a bargaining framework, in Gul (1989). In addition to the 'bargaining game' with which Gul gives a noncooperative foundation for the Shapley value, Gul describes a 'partnership game' which - in three player environments - implements the Shapley value of the homomollifier.

(We prove Proposition 5 in the Appendix. Proposition 5 can also be re-written with balanced consolidation-externalities in place of undominated consolidation-externalities.)

4. Externalities as 'threats'

One advantage shared by both the balanced contributions characterization of the Shapley value and the merge-externalities characterization of our Proposition 2, is that the axioms seem to translate intuitively to reference points ('threat points') to which players may realistically appeal while bargaining over the division of surplus. Balanced contributions can be interpreted as an equilibrium among 'threats' to simply walk away from negotiations. A violation of undominated merge-externalities can be interpreted as a disequilibrium among 'threats' to 'sell-up to', 'buy-out', or 'join forces with' other players. Whether one, or other, or both courses of action (walking away, or selling-up) are actually possible will of course depend on the specific situation of interest. Moreover, for the possibilities to actually be perceived, by other players, as 'threats', we might suppose that the external consequences should be negative. Whether the external consequences (of a player leaving, or of two players merging) are positive or negative will depend on the underlying game (N, v).

We aim to capture this idea - that only negative externalities constitute threats - by two possible further relaxations of the undominated merge-externalities condition. The first condition, *undominated merge-threats*, permits merge-externalities to be dominated if none of these externalities are negative. The second condition, *undominated-or-equivocal merge-threats*, goes further and permits merge-externalities to be dominated if any of these externalities are non-negative.

Undominated merge-threats. $\forall (N, v) \in \Gamma, |N| > 2 \rightarrow \forall i, j \in N$,
Either: $\exists k \in (N \setminus \{i, j\})$,
$\phi_i (\mathcal{M}((N, v), (j, k))) - \phi_i(N, v) \geqslant \phi_j (\mathcal{M}((N, v), (i, k))) - \phi_j(N, v))$.
Or: $\forall k \in (N \setminus \{i, j\}), \phi_i (\mathcal{M}((N, v), (j, k))) - \phi_i(N, v) \geqslant 0$.

Undominated-or-equivocal merge-threats.
$\forall (N, v) \in \Gamma, |N| > 2 \rightarrow \forall i, j \in N$,
Either: $\exists k \in (N \setminus \{i, j\})$,
$\phi_i (\mathcal{M}((N, v), (j, k))) - \phi_i(N, v) \geqslant \phi_j (\mathcal{M}((N, v), (i, k))) - \phi_j(N, v))$.
Or: $\exists k \in (N \setminus \{i, j\}), \phi_i (\mathcal{M}((N, v), (j, k))) - \phi_i(N, v) \geqslant 0$.

Following Gul (1989), we say that (N, v) is *value-additive* if and only if for any $\pi \in \Pi(N)$, and for any $I, J \in \pi$, $\phi^{Sh}_{I \cup J} \left(\overline{\mathcal{M}}((N, v)^{\pi}, \{I, J\})\right) \geqslant$

$\phi_I^{Sh}((N,v)^\pi) + \phi_J^{Sh}((N,v)^\pi)$. Following McQuillin and Sugden (2016) we say that there are *no positive value-externalities in* (N,v) if and only if for any $\pi \in \Pi(N)$, and for any $I, J, K \in \pi$, $\phi_I^{Sh}\left(\overline{\mathcal{M}}((N,v)^\pi, \{J,K\})\right) - \phi_I^{Sh}((N,v)^\pi) \leqslant 0$. We let $\widehat{\Gamma} \subset \Gamma$ denote the set of all games that are value-additive, and $\widetilde{\Gamma} \subset \widehat{\Gamma}$ the set of all games with no positive value-externalities.

Proposition 6. *(i) A solution ϕ satisfies, on the sub-class $\widehat{\Gamma}$, efficiency, equal-gains in 2-player games and undominated merge-threats, if and only if for any $(N,v) \in \widehat{\Gamma}$, $\phi(N,v) = \phi^{Sh}(N,v)$.*

(ii) A solution ϕ satisfies, on the sub-class $\widetilde{\Gamma}$, efficiency, equal-gains in 2-player games and undominated-or-equivocal merge-threats, if and only if for any $(N,v) \in \widetilde{\Gamma}$, $\phi(N,v) = \phi^{Sh}(N,v)$.

(We prove Proposition 6 in the Appendix.)

So, the Shapley value, postulated as the expected outcome of rational bargaining, may be seen as progressively more compelling when applied to games that are value-additive and to games with no positive value-externalities. This conclusion corresponds exactly with findings in analyses of noncooperative bargaining models - specifically those of Gul (1989) and McQuillin and Sugden (2016) - that allow for bilateral mergers. In Gul's bargaining model, value-additivity is a necessary and sufficient condition for the existence of a stationary subgame perfect equilibrium within which players' expectations converge (as a discount factor tends to 1) to the Shapley value, and (McQuillin and Sugden show) no positive value-externalities ensures that there are no other stationary subgame perfect equilibria. In McQuillin and Sugden's model, no positive value-externalities ensures that every subgame perfect equilibrium supports the Shapley value. It would be useful to understand better which types of games have these properties: Haller (1994) and Derks and Tijs (2000) make some progress in this direction.

5. Monotonic simple games

A further advantage shared by both the balanced contributions characterization of the Shapley value and the merge-externalities characterization of our Proposition 2 comes from the fact that the characterizations also hold on useful restrictions ('sub-classes') of Γ. For example the balanced contributions characterization is unusual (among the well-known characterizations of the Shapley value) in that it holds on the class of 'assignment games' (defined by Shapley and Shubik, 1971).[12] If (N,v) is an assignment game then

[12] Brink and Pintér (2015) provide an alternative to the Shapley value solution that satisifes all of the axioms in many of the best-known characterizations of the Shapley

so will be any leave subgame of (N, v) (but a merge subgame may not be).

On the other hand, our 'merge-externalities' characterization is unusual in that it holds on the class of 'monotonic simple games'. (We say that $(N, v) \in \Gamma$ is a 'simple game' if it has properties (i) $v(N) = 1$, and (ii) for any $S \subset N$, $v(S) \in \{0, 1\}$. And a simple game (N, v) is 'monotonic' if for any $S \subset N$ and for any $I \subset S$, $v(I) = 1 \to v(S) = 1$.) Shapley and Shubik (1954) proposed using monotonic simple games to model voting rights, and using the Shapley value in this context as a measure of *a priori voting power*. This measure is now known as the *Shapley-Shubik power index*. It is well known that the original Shapley axioms do not characterize a unique solution on the sub-class of such games[13], but it is very straightforward to see that if (N, v) is a simple game then any merge subgame of (N, v) will be a simple game (but a leave subgame may not be), and also that if (N, v) is monotonic then any merge subgame of (N, v) will be monotonic. So therefore the Shapley-Shubik power index can also be characterized by efficiency, equal-gains in 2-player games and undominated merge-externalities. In addition, if a simple game (N, v) is *strong* ($\forall S \subseteq N, v(S) = 0 \to v(N \setminus S) = 1$) or *proper* ($\forall S \subseteq N, v(S) = 1 \to v(N \setminus S) = 0$) then any merge subgame will also be a strong or a proper simple game respectively; so the characterization holds similarly on these further sub-classes.

6. Balanced externalities and the *extended* Shapley value

The Shapley value represents an important proposition about the outcome of unstructured bargaining among rational agents, applicable to bargaining problems that are adequately represented as games in characteristic function form. But how the value should be extended to games in *partition function* form (as first described in Thrall and Lucas, 1963) remains an open question. (A number of alternative proposals have been advanced, but there is no present consensus on their relative applicability to specific problems.) This issue has important implications because many real-world bargaining problems are naturally represented as games in partition function form – that is, as cooperative games in which the payoff to one coalition of players can depend on how the other players are grouped into coalitions. For example,

value, once these axioms are restricted to hold on assignment games only. They also show that conditions used by Myerson (1977) - 'component efficiency' and 'fairness' - can be used to characterize the Shapley value on the class of assignment games.

[13] Dubey (1975) provided the first characterization of the Shapley value for monotonic simple games (and therefore of the Shapley-Shubik power index), with an adaptation of the additivity axiom (now known as the 'transfer' axiom) to operate within this sub-class.

consider a set of firms which have opportunities to merge or to coalesce into price-fixing cartels. The payoff to a given coalition of merged or colluding firms typically depends on the pattern of coalescence among the other firms. Alternatively, consider the circumstance of nations negotiating environmental agreements: the eventual well-being of any nation is affected both by agreements it joins and by agreements, between other nations, it does not join.

A difficulty arises when one tries to extend the balanced contributions approach to games in partition function form. If the original game is in characteristic function form then the subgames that arise as players leave are unambiguous, whereas if the original game is in partition function form then the new games that arise are unclear. In particular, if several players leave, we cannot tell whether the players that remain should suppose that the absent players are organized as singletons, coalesced together, or in some other configuration. One can produce a characterization that employs one or other supposition (if we suppose that the absent players remain as singletons then this leads to the value proposed by Pham Do and Norde, 2007, and by de Clippel and Serrano, 2008) but the supposition seems arbitrary. (The ideas of applying the 'balanced contributions' approach to games in partition function form, and of this associated difficulty appear in de Clippel and Serrano, 2008.)

Our merge-externalities approach extends without this difficulty, because even after multiple mergers the relevant partition of the original player set seems clear. (If i merges with j, and then if the new merged entity proceeds to merge with k, it then seems clear that any remaining players should now view i, j and k as a coalition.) It turns out that the merge-externalities approach supports the value proposed by McQuillin (2009), the distinctive feature of which is that players' expectations are independent of payoffs in the underlying game associated with partitions of cardinality greater than two. It should be emphasized that this feature emerges as a result (not as an assumption) both within the axiomatic approach of McQuillin (2009) and, for a defined sub-class of games (games that have 'no positive value-externalities'), within the noncooperative bargaining model of McQuillin and Sugden (2016).

An intuition for the claim that some payoffs in an underlying game may be nugatory to the bargaining outcome can be gained by considering a modified version of the 3-player glove game discussed in the introduction. Suppose that the payoff for one player, $i \in \{1, 2, 3\}$, as a singleton in this game, is now modified to $x > 0$ in the circumstance that the other two players remain as singletons. (We suppose that i's payoff reverts to zero if the two other players coalesce.) Is i's bargaining position - in an environment such that players are able to merge and then continue to negotiate - strengthened in the modified bargaining game, relative to the original? It is hard to see how. In the modified game, just as in the original, the other two players can strengthen their collective position by merging. To avert this, i must herself seek out some merger. The singleton structure, in which player i receives x, does not appear as the natural counterfactual to any of these possible transactions, and is therefore not considered in negotiations. Our remaining formal results generalize this reasoning.

Given any set of players N, let $M(N)$ denote the set of *embedded coalitions*, $M(N) \equiv \{(S,\pi) : \pi \in \Pi(N), S \in \pi\}$. Let Γ^+ denote the set of *transferable utility games in partition function form*, generically (N,w) with $w\colon M(N) \to \mathbb{R}$ having the property for any $\pi \in N$, $w(\varnothing,\pi) = 0$. We shall now suppose that a solution, generically φ, associates with each $(N,w) \in \Gamma^+$ and each $i \in N$ a real number $\varphi_i(N,w)$, representing i's utility expectation in the game (N,w).

One *extended Shapley value*, due to Pham Do and Norde (2007) and de Clippel and Serrano (2008), φ^{ESV1}, is defined using (for any $(N,w) \in \Gamma^+$, and for any $i \in N$):

$$\begin{aligned} \varphi_i^{ESV1}(N,w) &= \phi_i^{Sh}(N,v) : \\ \forall S &\subseteq N, v(S) = w(S, \{S\} \cup \{\{j\}_{j \in (N \setminus S)}\}). \end{aligned}$$

Another extended Shapley value, due to McQuillin (2009), φ^{ESV2}, is defined using (for any $(N,w) \in \Gamma^+$, and for any $i \in N$):

$$\begin{aligned} \varphi_i^{ESV2}(N,w) &= \phi_i^{Sh}(N,v) : \\ \forall S &\subseteq N, v(S) = w(S, \{S, N \setminus S\}). \end{aligned}$$

We additionally define, for any $(N,w) \in \Gamma^+$, and for any $i,j \in N$, subgames $\mathcal{L}((N,w),i) \in \Gamma^+$ and $\mathcal{M}((N,w),(i,j)) \in \Gamma^+$, using:

$$\begin{aligned} \mathcal{L}((N,w),i) &\equiv (N \setminus \{i\}, w') : \\ \forall (S,\pi) &\in M(N \setminus \{i\}), w'(S,\pi) = w(S, \pi \cup \{\{i\}\}). \end{aligned}$$

$$\begin{aligned} \mathcal{M}((N,w),(i,j)) &\equiv (N \setminus \{i\}, w') : \forall (S,\pi) \in M(N \setminus \{i\}), \\ w'(S,\pi) &= \begin{cases} w(S, \{I\}_{I \in \pi : j \notin I} \cup \{I \cup \{i\}\}_{I \in \pi : j \in I}) & \text{where } j \notin S \\ w(S \cup \{i\}, \{I\}_{I \in \pi \setminus \{S\}} \cup \{S \cup \{i\}\}) & \text{where } j \in S \end{cases}. \end{aligned}$$

The interpretations of $\mathcal{L}((N,w),i)$ and $\mathcal{M}((N,w),(i,j))$ remain, as before, subgames formed respectively by i 'leaving' or i 'merging with j'.

As we have already noted, in describing subgames that arise from multiple players successively *leaving* we are forced to make an assumption about how the departed players are partitioned. For example, consider a game $(\{i,j,k,l\},w) \in \Gamma^+$, from which players k and l consecutively leave, generating a subgame $(\{i,j\},w')$. Here, we have supposed that the players who leave the game remain as singletons, so w' is defined $w'(\{i\},\{\{i\},\{j\}\}) = w(\{i\},\{\{i\},\{j\},\{k\},\{l\}\})$, $w'(\{j\},\{\{i\},\{j\}\}) = w(\{j\},\{\{i\},\{j\},\{k\},\{l\}\})$, $w'(\{i,j\},\{\{i,j\}\}) = w(\{i,j\},\{\{i,j\},\{k\},\{l\}\})$. But, alternatively, we *could* have supposed that players who leave all coalesce together - giving $w'(\{i\},\{\{i\},\{j\}\}) = w(\{i\},\{\{i\},\{j\},\{k,l\}\})$, and so forth - or indeed that the absent players in some sense randomize over possible structures.

In describing subgames that arise from players *merging* there is no similar difficulty. Each amalgamation leads to a well-defined partition of the original player set. For example, if the subgame $(\{i,j\},w')$ is formed from $(\{i,j,k,l\},w)$ by players k and l consecutively merging with j, then our conception of merger itself directly entails $w'(\{j\},\{\{i\},\{j\}\}) = w(\{j,k,l\},\{\{i\},\{j,k,l\}\})$ and $w'(\{i,j\},\{\{i,j\}\}) = w(\{i,j,k,l\},\{\{i,j,k,l\}\})$. If $\{j\}$ is now acting as the coalition $\{j,k,l\}$, then the payoff to $\{i\}$ should be as previously specified when $\{j,k,l\}$ has formed, so we must also have $w'(\{i\},\{\{i\},\{j\}\}) = w(\{i\},\{\{i\},\{j,k,l\}\})$.[14]

We now re-define the conditions we previously proposed, this time for a solution φ on games in Γ^+.

Efficiency$^+$. $\forall (N,w) \in \Gamma^+, \sum_{i \in N} \varphi_i(N,w) = w(N,\{N\})$.

Equal-gains in 2-player games$^+$. $\forall (N,w) \in \Gamma^+, |N| = 2 \rightarrow \forall i,j \in N, \varphi_i(N,w) - w(\{i\},\{\{i\},\{j\}\}) = \varphi_j(N,w) - w(\{j\},\{\{i\},\{j\}\})$.

Balanced contributions$^+$. $\forall (N,w) \in \Gamma^+, \forall i,j \in N, \varphi_i(\mathcal{L}((N,w),j)) - \varphi_i(N,w) = \varphi_j(\mathcal{L}((N,w),i)) - \varphi_j(N,w)$.

Balanced merge-externalities$^+$. $\forall (N,w) \in \Gamma^+, \forall i,j,k \in N$,
$\varphi_i(\mathcal{M}((N,w),(j,k))) - \varphi_i(N,w) = \varphi_j(\mathcal{M}((N,w),(i,k))) - \varphi_j(N,w)$.

Undominated merge-externalities$^+$. $\forall (N,w) \in \Gamma^+, |N| > 2 \rightarrow \forall i,j \in N, \exists k \in N \setminus \{i,j\}$,
$\varphi_i(\mathcal{M}((N,w),(j,k))) - \varphi_i(N,w) \geqslant \varphi_j(\mathcal{M}((N,w),(i,k))) - \varphi_j(N,w)$.

[14] The idea that each amalgamation leads to a well-defined partition of the original player set is even clearer within the 'consolidation' conception of merger used above in the final characterizations of Section 3; however, for the present section, we retain the notationally simpler conception used elsewhere in the paper.

De Clippel and Serrano (2008) have already noted that efficiency$^+$ and balanced contributions$^+$ together entail $\varphi = \varphi^{ESV1}$. It turns out that efficiency$^+$, equal-gains in 2-player games$^+$, and undominated merge-externalities$^+$ together entail $\varphi = \varphi^{ESV2}$.

Proposition 7. *A solution φ satisfies efficiency$^+$, equal-gains in 2-player games$^+$ and undominated merge-externalities$^+$ if and only if $\varphi = \varphi^{ESV2}$. Balanced merge-externalities$^+$ is also satisfied by $\varphi = \varphi^{ESV2}$.*

(We prove Proposition 6 in the Appendix.)

The analogous results to those presented in Section 4 also hold. For any $(N, w) \in \Gamma^+$, we define $\Gamma^+_{(N,w)}$ to be the set of all subgames, including (N, w), reachable by sequences of mergers from (N, w). We say that (N, w) is *value-additive* if and only if for any $(N', w') \in \Gamma^+_{(N,w)}$, and for any $i, j \in N'$, $\varphi^{ESV2}_j (\mathcal{M}((N', w'), (i, j))) \geqslant \varphi^{ESV2}_i(N', w') + \varphi^{ESV2}_j(N', w')$. We say that there are *no positive value-externalities in* (N, w) if and only if for any $(N', w') \in \Gamma^+_{(N,w)}$, and for any $i, j, k \in N'$, $\varphi^{ESV2}_i (\mathcal{M}((N', w'), (j, k))) - \varphi^{ESV2}_i(N', w') \leqslant 0$. We let $\widehat{\Gamma^+} \subset \Gamma^+$ denote the set of all partition function games that are value-additive, and $\widetilde{\Gamma^+} \subset \widehat{\Gamma^+}$ the set of all such games with no positive value-externalities.

Undominated merge-threats$^+$. $\forall (N, w) \in \Gamma^+, |N| > 2 \rightarrow \forall i, j \in N,$
Either: $\exists k \in (N \setminus \{i, j\})$,
$\varphi_i (\mathcal{M}((N, w), (j, k))) - \varphi_i(N, w) \geqslant \varphi_j (\mathcal{M}((N, w), (i, k))) - \varphi_j(N, w))$.
Or: $\forall k \in (N \setminus \{i, j\}), \varphi_i (\mathcal{M}((N, w), (j, k))) - \varphi_i(N, w) \geqslant 0$.

Undominated-or-equivocal merge-threats$^+$. $\forall (N, w) \in \Gamma^+, |N| > 2 \rightarrow \forall i, j \in N,$
Either: $\exists k \in (N \setminus \{i, j\})$,
$\varphi_i (\mathcal{M}((N, w), (j, k))) - \varphi_i(N, v) \geqslant \varphi_j (\mathcal{M}((N, w), (i, k\})) - \varphi_j(N, w))$.
Or: $\exists k \in (N \setminus \{i, j\}), \varphi_i (\mathcal{M}((N, w), (j, k))) - \varphi_i(N, w) \geqslant 0$.

Proposition 8. *(i) A solution φ satisfies, on the sub-class $\widehat{\Gamma^+}$, efficiency$^+$, equal-gains in 2-player games$^+$ and undominated merge-threats$^+$, if and only if for any $(N, w) \in \widehat{\Gamma^+}$, $\varphi (N, w) = \varphi^{ESV2} (N, w)$.*

(ii) A solution φ satisfies, on the sub-class $\widetilde{\Gamma^+}$, efficiency$^+$, equal-gains in 2-player games$^+$ and undominated-or-equivocal merge-threats$^+$, if and only if for any $(N, w) \in \widetilde{\Gamma^+}$, $\varphi (N, w) = \varphi^{ESV2} (N, w)$.

By Proposition 7, within our merge-externalities approach, the solution φ^{ESV2} emerges as the analogue, for games in partition function form, to the solution ϕ^{Sh} for games in characteristic function form. By Proposition 8, and consistent with the result in McQuillin and Sugden (2016), this solution is most defensible as the expected outcome from rational bargaining for games (such as the glove game) in which there are 'no positive value-externalities'.

Acknowledgements

We thank the Special Issue editor and two anonymous referees for their comments. Sugden's work was supported by the Economic and Social Research Council of the UK (award no. ES/K002201/1).

Appendix

Proofs of Propositions 1-8

Proof of Proposition 1. For any semivalue ϕ, for any $(N,v) \in \Gamma$, for any $i,j,k \in N$, there are probability measures $\{p_s^n : n \in \mathbb{Z}_+, s \in \{0,...,n-1\}\}$ satisfying $\sum_{t=0}^{n-1} \binom{n-1}{t} p_t^n = 1$ and, due to $t^{s+1}(1-t)^{n-(s+1)-1} + t^s(1-t)^{n-s-1} = t^s(1-t)^{(n-1)-s-1}$, $p_{s+1}^n + p_s^n = p_s^{n-1}$ such that we can write:

$$\begin{aligned}
&\phi_i\left(\mathcal{M}((N,v),(j,k))\right) - \phi_i(N,v) \\
&\quad = \sum_{\substack{S\subseteq N\setminus\{i\}: \\ \{j,k\}\cap S=\emptyset}} p_{|S|}^{|N|-1}\left(v(S\cup\{i\}) - v(S)\right) \\
&\qquad + \sum_{\substack{S\subseteq N\setminus\{i\}: \\ \{j,k\}\subseteq S}} p_{|S|-1}^{|N|-1}\left(v(S\cup\{i\}) - v(S)\right) \\
&\qquad - \sum_{S\subseteq N\setminus\{i\}} p_{|S|}^{|N|}\left(v(S\cup\{i\}) - v(S)\right). \qquad (.1)
\end{aligned}$$

Rearranging the right hand side of (.1) gives:

$$
\begin{aligned}
&\phi_i\left(\mathcal{M}((N,v),(j,k))\right) - \phi_i(N,v)\\
&= \sum_{S\subseteq (N\setminus\{i,j,k\})} \Bigg(\left(p_{|S|}^{|N|-1} v(S\cup\{i\}) - p_{|S|}^{|N|-1} v\left(S\right)\right)\\
&\qquad + \left(p_{|S|+1}^{|N|-1} v(S\cup\{i,j,k\}) - p_{|S|+1}^{|N|-1} v\left(S\cup\{j,k\}\right)\right)\\
&\qquad - \left(p_{|S|}^{|N|} v(S\cup\{i\}) - p_{|S|}^{|N|} v\left(S\right)\right)\\
&\qquad - \left(p_{|S|+1}^{|N|} v(S\cup\{i,j\}) - p_{|S|+1}^{|N|} v\left(S\cup\{j\}\right)\right)\\
&\qquad - \left(p_{|S|+1}^{|N|} v(S\cup\{i,k\}) - p_{|S|+1}^{|N|} v\left(S\cup\{k\}\right)\right)\\
&\qquad - \left(p_{|S|+2}^{|N|} v(S\cup\{i,j,k\}) - p_{|S|+2}^{|N|} v\left(S\cup\{j,k\}\right)\right)\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{S\subseteq (N\setminus\{i,j,k\})} \Bigg(\left(p_{|S|}^{|N|} - p_{|S|}^{|N|-1}\right) v(S) + \left(p_{|S|}^{|N|-1} - p_{|S|}^{|N|}\right) v(S\cup\{i\})\\
&\quad + p_{|S|+1}^{|N|} v\left(S\cup\{j\}\right) + p_{|S|+1}^{|N|} v\left(S\cup\{k\}\right) - p_{|S|+1}^{|N|} v(S\cup\{i,j\})\\
&\quad - p_{|S|+1}^{|N|} v(S\cup\{i,k\}) + \left(p_{|S|+2}^{|N|} - p_{|S|+1}^{|N|-1}\right) v(S\cup\{j,k\})\\
&\quad - p_{|S|+2}^{|N|} v(S\cup\{i,j,k\})\Bigg)\\
&= \sum_{S\subseteq (N\setminus\{i,j,k\})} \Big(-p_{|S|+1}^{|N|} v(S) + p_{|S|+1}^{|N|} v(S\cup\{i\}) + p_{|S|+1}^{|N|} v\left(S\cup\{j\}\right)\\
&\quad + p_{|S|+1}^{|N|} v\left(S\cup\{k\}\right) - p_{|S|+1}^{|N|} v(S\cup\{i,j\}) - p_{|S|+1}^{|N|} v(S\cup\{i,k\})\\
&\quad - p_{|S|+1}^{|N|} v(S\cup\{j,k\}) - p_{|S|+2}^{|N|} v(S\cup\{i,j,k\})\Big). \qquad (.2)
\end{aligned}
$$

So, by symmetries in the coefficients on the right hand side of (.2), $\phi_i\left(\mathcal{M}((N,v)\,,(j,k))\right) - \phi_i(N,v) = \phi_j\left(\mathcal{M}((N,v),(i,k))\right) - \phi_j(N,v)$. ■

Proof of Proposition 2. It is immediately obvious that $\phi = \phi^{Sh}$ satisfies efficiency and equal-gains in 2-player games. Also, ϕ^{Sh} is a semivalue, so by Proposition 1 $\phi = \phi^{Sh}$ satisfies balanced (and therefore undominated) merge-externalities. It remains to show that if a solution $\phi \neq \phi^{Sh}$ satisfies efficiency and equal-gains in 2-player games then ϕ contravenes undominated merge-externalities. Note that if ϕ satisfies efficiency and equal-gains in 2-player games then $\forall (N,v) \in \Gamma, |N| \leqslant 2 \rightarrow \phi(N,v) = \phi^{Sh}(N,v)$. Suppose $\phi \neq \phi^{Sh}$ satisfies efficiency and equal-gains in 2-player games and select the highest positive integer n such that $\forall (N,v) \in \Gamma, |N| < n \rightarrow \phi(N,v) = \phi^{Sh}(N,v)$.

Consider any $(N, v) \in \Gamma$ such that $|N| = n$ and $\phi(N, v) \neq \phi^{Sh}(N, v)$. Then select $i \in N, j \in N$ such that $\phi_i(N, v) > \phi_i^{Sh}(N, v), \phi_j(N, v) < \phi_j^{Sh}(N, v)$. (We know, through our selection of (N, v) and because ϕ and ϕ^{Sh} both satisfy efficiency, that some such i and j exist.) For any $k \in N \setminus \{i, j\}$, $\phi_i^{Sh}(\mathcal{M}((N, v), (j, k))) - \phi_i^{Sh}(N, v) = \phi_j^{Sh}(\mathcal{M}((N, v), (i, k))) - \phi_j^{Sh}(N, v)$, also $\phi_i(\mathcal{M}((N, v), (j, k))) = \phi_i^{Sh}(\mathcal{M}((N, v), (j, k)))$ and $\phi_j(\mathcal{M}((N, v), (i, k))) = \phi_j^{Sh}(\mathcal{M}((N, w), (i, k)))$, therefore $\phi_i(\mathcal{M}((N, v), (j, k))) - \phi_i(N, v) < \phi_j(\mathcal{M}((N, v), (i, k))) - \phi_j(N, v)$. ■

Proof of Proposition 3. It is immediately obvious that $\phi = \phi^{Bz}$ satisfies equal-gains in 2-player games and Bz-sum. Also, ϕ^{Bz} is a semivalue, so by Proposition 1 $\phi = \phi^{Bz}$ satisfies balanced (and therefore undominated) merge-externalities. The remainder of the proof (showing that if a solution $\phi \neq \phi^{Bz}$ satisfies equal-gains in 2-player games and Bz-sum then ϕ contravenes undominated merge-externalities) simply replicates the inductive argument in the proof above of Proposition 2. ■

Proof of Proposition 4. It is immediately obvious that $\phi = \phi^{ED}$ satisfies efficiency and equal-division in 2-player games. Also, we have $\phi_i^{ED}(\mathcal{M}((N, v), (j, k))) - \phi_i^{ED}(N, v) = \frac{v(N)}{|N|-1} - \frac{v(N)}{|N|} = \phi_j^{ED}(\mathcal{M}((N, v), (i, k))) - \phi_j^{ED}(N, v)$, so $\phi = \phi^{ED}$ satisfies balanced merge-externalities. The remainder of the proof (showing that if a solution $\phi \neq \phi^{ED}$ satisfies efficiency and equal-division in 2-player games then ϕ contravenes undominated merge-externalities) replicates the inductive arguments in the proofs above. ■

Proof of Proposition 5(i).

Note that the sub-class $Q(N, v)$ contains all of its merge-subgames. It is immediately obvious that if $\forall (N, v)^\pi \in Q(N, v), \phi((N, v)^\pi) = \phi^{Sh}((N, v)^\pi)$ then ϕ satisfies efficiency and equal-gains in 2-player games on the sub-class $Q(N, v)$. Also (using the definition of ϕ^{Sh}, and using the balanced merge-externalities property of ϕ^{Sh} established in Proposition 1), $\forall (\pi, v^\pi) \in Q(N, v), \forall I, J, K \in \pi$:

$$\begin{aligned}
&\phi_I^{Sh}\left(\overline{\mathcal{M}}((\pi, v^\pi), \{J, K\})\right) - \phi_I^{Sh}(\pi, v^\pi) \\
&\quad = \phi_I^{Sh}(\mathcal{M}((\pi, v^\pi), (J, K))) - \phi_I^{Sh}(\pi, v^\pi) \\
&\quad = \phi_J^{Sh}(\mathcal{M}((\pi, v^\pi), (I, K))) - \phi_J^{Sh}(\pi, v^\pi) \\
&\quad = \phi_J^{Sh}\left(\overline{\mathcal{M}}((\pi, v^\pi), \{I, K\})\right) - \phi_J^{Sh}(\pi, v^\pi).
\end{aligned}$$

So if $\forall (N, v)^\pi \in Q(N, v), \phi((N, v)^\pi) = \phi^{Sh}((N, v)^\pi)$ then ϕ satisfies balanced consolidation-externalities on the sub-class $Q(N, v)$. The remainder of the proof (showing that if a solution ϕ such that $\exists (N, v)^\pi \in Q(N, v)$, $\phi((N, v)^\pi) \neq \phi^{Sh}((N, v)^\pi)$ satisfies efficiency and equal-gains in 2-player games on the sub-class $Q(N, v)$ then ϕ contravenes undominated merge-externalities on the sub-class $Q(N,v)$) replicates the inductive arguments in the proofs above. ■

Proof of Proposition 5(ii). First notice that $h_v(N) = v(N)$, so if $\forall (N,v)^\pi \in Q(N,v), \phi((N,v)^\pi) = \phi^{Sh}((N,h_v)^\pi)$ then ϕ (by the efficiency property of ϕ^{Sh}) satisfies efficiency on the sub-class $Q(N,v)$; also (by the same argument used in proof of Proposition 5, substituting h_v for v), ϕ satisfies balanced consolidation-externalities on the sub-class $Q(N,v)$. Using the definitions of ϕ^{Sh} and of h_v:

$$\begin{aligned}
\forall (N,v)^\pi \in Q(N,v), |\pi| = 2 \quad &\rightarrow \quad \forall I, J \in \pi, \\
\phi_I^{Sh}((N,h_v)^\pi) \quad &= \quad \tfrac{1}{2}\left(\tfrac{|I|(v(N)-v(J))+|J|v(I)}{|N|}\right) \\
&\quad\; + \tfrac{1}{2}\left(v(N) - \left(\tfrac{|J|(v(N)-v(I))+|I|v(J)}{|N|}\right)\right) \\
&= \quad v(I) + \tfrac{|I|}{|N|}(v(N) - v(I) - v(J)) \\
\phi_J^{Sh}((N,h_v)^\pi) \quad &= \quad v(J) + \tfrac{|J|}{|N|}(v(N) - v(I) - v(J)).
\end{aligned}$$

So if $\forall (N,v)^\pi \in Q(N,v), \phi((N,v)^\pi) = \phi^{Sh}((N,h_v)^\pi)$ then ϕ also satisfies equal per-capita gains in 2-player games. The remainder of the proof (showing that if a solution ϕ such that $\exists (N,v)^\pi \in Q(N,v), \phi((N,v)^\pi) \neq \phi^{Sh}((N,h_v)^\pi)$ satisfies efficiency and equal per-capita gains in 2-player games on the sub-class $Q(N,v)$ then ϕ contravenes undominated merge-externalities on the sub-class $Q(N,v)$) replicates the inductive arguments in the proofs above. ■

Proof of Proposition 6(i). Note that the sub-class $\widehat{\Gamma}$ contains all of its merge-subgames. By Proposition 2, if $\forall (N,v) \in \widehat{\Gamma}, \phi(N,v) = \phi^{Sh}(N,v)$ then ϕ satisfies, on the sub-class $\widehat{\Gamma}$, efficiency, equal-gains in 2-player games and undominated merge-externalities (therefore also undominated merge-threats). Now assume ϕ such that $\exists (N,v) \in \widehat{\Gamma}, \phi(N,v) \neq \phi^{Sh}(N,v)$ satisfies efficiency and equal-gains in 2-player games on the sub-class $\widehat{\Gamma}$. Note that $\forall (N,v) \in \widehat{\Gamma}, |N| \leqslant 2 \rightarrow \phi(N,v) = \phi^{Sh}(N,v)$. Select the highest positive integer n such that $\forall (N,v) \in \widehat{\Gamma}, |N| < n \rightarrow \phi(N,v) = \phi^{Sh}(N,v)$. Consider any $(N,v) \in \widehat{\Gamma}$ such that $|N| = n$ and $\phi(N,v) \neq \phi^{Sh}(N,v)$. Then select $i \in N, j \in N$ such that $\phi_i(N,v) > \phi_i^{Sh}(N,v), \phi_j(N,v) < \phi_j^{Sh}(N,v)$. (We know, through our selection of (N,v) and because ϕ and ϕ^{Sh} both satisfy efficiency on $\widehat{\Gamma}$, that some such i and j exist.) For any $k \in N \setminus \{i,j\}$, by the balanced merge-externalities property of ϕ^{Sh} established in Proposition 1, $\phi_i^{Sh}(\mathcal{M}((N,v), (j,k))) - \phi_i^{Sh}(N,v) = \phi_j^{Sh}(\mathcal{M}((N,v), (i,k))) - \phi_j^{Sh}(N,v)$, also $\phi_i(\mathcal{M}((N,v), (j,k))) = \phi_i^{Sh}(\mathcal{M}((N,v), (j,k)))$ and $\phi_j(\mathcal{M}((N,v), (i,k))) = \phi_j^{Sh}(\mathcal{M}((N,v), (i,k)))$, therefore $\phi_i(\mathcal{M}((N,v), (j,k))) - \phi_i(N,v) < \phi_j(\mathcal{M}((N,v), (i,k))) -$

$\phi_j(N,v)$. Now select $k \in N \setminus \{i,j\}$ such that $\phi_i^{Sh}(\mathcal{M}((N,v),(j,k))) - \phi_i^{Sh}(N,v) \leqslant 0$. (To see that some such k exists, note that by value-additivity of (N,v), $\phi_j^{Sh}(\mathcal{M}((N,v),(i,j))) \geqslant \phi_i^{Sh}(N,v) + \phi_j^{Sh}(N,v)$. Therefore, $\exists k \in N\setminus\{i,j\}, \phi_k^{Sh}(\mathcal{M}((N,v),(i,j))) - \phi_k^{Sh}(N,v) \leqslant 0$. Clearly, $\phi_i^{Sh}(\mathcal{M}((N,v),(j,k))) = \phi_i^{Sh}(\mathcal{M}((N,v),(k,j)))$ and moreover - again by Proposition 1 - we have: $\phi_i^{Sh}(\mathcal{M}((N,v),(k,j))) - \phi_i^{Sh}(N,v) = \phi_k^{Sh}(\mathcal{M}((N,v),(i,j))) - \phi_k^{Sh}(N,v)$. Therefore, $\exists k \in N \setminus \{i,j\}, \phi_i^{Sh}(\mathcal{M}((N,v),(j,k))) - \phi_i^{Sh}(N,v) \leqslant 0$.) Then $\phi_i(\mathcal{M}((N,v),(j,k))) - \phi_i(N,v) < \phi_i^{Sh}(\mathcal{M}((N,v),(j,k))) - \phi_i^{Sh}(N,v) \leqslant 0$. So ϕ contravenes undominated merge-threats on the sub-class $\widehat{\Gamma}$. ■

Proof of Proposition 6(ii). We repeat the arguments above, replacing $\widehat{\Gamma}$ with $\widetilde{\Gamma}$. At the last step, note that (because $(N,v) \in \widetilde{\Gamma}$ has no positive value-externalities) $\forall k \in N \setminus \{i,j\}, \phi_i^{Sh}(\mathcal{M}((N,v),(j,k))) - \phi_i^{Sh}(N,v) \leqslant 0$. Therefore, $\forall k \in N \setminus \{i,j\}, \phi_i(\mathcal{M}((N,v),(j,k))) - \phi_i(N,v) < 0$. So ϕ contravenes undominated-or-equivocal merge-threats on the sub-class $\widetilde{\Gamma}$. ■

Proof of Proposition 7. It is immediately obvious that $\varphi = \varphi^{ESV2}$ satisfies efficiency$^+$ and equal-gains in 2-player games$^+$. For any $(N,w) \in \Gamma^+$ we define $(N,v^w) \in \Gamma$ using: $\forall S \subseteq N, v^w(S) = w(S,\{S, N \setminus S\})$. Clearly, $v^{\mathcal{M}((N,w),(j,k))} = \mathcal{M}((N,v^w),(j,k))$, and therefore $\varphi_i^{ESV2}(\mathcal{M}((N,w),(j,k))) - \varphi_i^{ESV2}(N,w) = \phi_i^{Sh}(\mathcal{M}((N,v^w),(j,k))) - \phi_i^{Sh}(N,v^w)$. So, since $\phi = \phi^{Sh}$ satisfies balanced merge-externalities, $\varphi = \varphi^{ESV2}$ satisfies balanced merge-externalities$^+$. It then remains to show that if a solution $\varphi \neq \varphi^{ESV2}$ satisfies efficiency$^+$ and equal-gains in 2-player games$^+$ then φ contravenes undominated merge-externalities$^+$. Note that if φ satisfies efficiency$^+$ and equal-gains in 2-player games$^+$ then $\forall (N,w) \in \Gamma^+, |N| \leqslant 2 \rightarrow \varphi(N,w) = \varphi^{ESV2}(N,w)$. Suppose $\varphi \neq \varphi^{ESV2}$ satisfies efficiency$^+$ and equal-gains in 2-player games$^+$ and select the highest positive integer n such that $\forall (N,w) \in \Gamma^+, |N| < n \rightarrow \varphi(N,w) = \varphi^{ESV2}(N,w)$. Consider any $(N,w) \in \Gamma^+$ such that $|N| = n$ and $\varphi(N,w) \neq \varphi^{ESV2}(N,w)$. Then select $i \in N, j \in N$ such that $\varphi_i(N,w) > \varphi_i^{ESV2}(N,w), \varphi_j(N,w) < \varphi_j^{ESV2}(N,w)$. (We know, through our selection of (N,w) and because φ and φ^{ESV2} both satisfy efficiency$^+$, that some such i and j exist.) For any $k \in N \setminus \{i,j\}$, $\varphi_i^{ESV2}(\mathcal{M}((N,w),(j,k))) - \varphi_i^{ESV2}(N,w) = \varphi_j^{ESV2}(\mathcal{M}((N,w),(i,k))) - \varphi_j^{ESV2}(N,w)$, also $\varphi_i(\mathcal{M}((N,w),(j,k))) = \varphi_i^{ESV2}(\mathcal{M}((N,w),(j,k)))$ and $\varphi_j(\mathcal{M}((N,w),(i,k))) = \varphi_j^{ESV2}(\mathcal{M}((N,w),(i,k)))$, therefore $\varphi_i(\mathcal{M}((N,w),(j,k))) - \varphi_i(N,w) < \varphi_j(\mathcal{M}((N,w),(i,k))) - \varphi_j(N,w)$. ■

Proofs of Propositions 8(i) and (ii). Proofs of Propositions 8(i) and (ii) are a matter only of substituting the relevant notation and condition names respectively in the proofs of Propositions 6(i) and (ii). ■

Buchanan JM (1987) The constitution of economic policy. Amer. Econ. Rev. 77, 243-50.

Binmore K (1998) Just Playing: Game Theory and the Social Contract Vol. 2, MIT Press.

van den Brink R, Funaki Y (2009) Axiomatizations of a class of equal surplus sharing solutions for TU-games. Theor. Decis 67, 303-340.

van den Brink R, Pintér M (2015) On axiomatizations of the Shapley value for assignment games. J. Math. Econ. 60, 110-114.

Casajus, A (2012) Amalgamating players, symmetry, and the Banzhaf value. Int J Game Theory 41, 497-515.

Charnes A, Rousseau J, Seiford L, (1978) Complements, mollifiers and the propensity to disrupt. Int. J. Game Theory 7, 37-50.

de Clippel G, Serrano R (2008) Marginal contributions and externalities in the value. Econometrica 76, 1413-1436.

Derks J, Tijs S (2000) On merge properties of the Shapley value. Int. Game Theor. Rev. 4, 249-257.

Dubey P (1975) On the uniqueness of the Shapley value. Int. J. Game Theory 4, 131-139.

Dubey P, Neyman A, Weber RJ (1981) Value theory without efficiency. Math. Oper. Res. 6, 122-128.

Dubey P, Shapley LS (1979) Mathematical properties of the Banzhaf power index. Math. Oper. Res. 4, 99-131.

Gul F (1989) Bargaining foundations of Shapley value, Econometrica 57, 81-95.

Haller H (1994) Collusion properties of values. Int. J. Game Theory 23, 261-281.

Hart S, Mas-Colell A (1989) Potential, value and consistency. Econometrica 57, 589-614.

Hart S, Mas-Colell A (1996) Bargaining and value. Econometrica 64, 357-380.

Lehrer E (1988) An axiomatization of the Banzhef value. Int. J. Game Theory 17, 89-99.

Malawski M (2002) Equal treatment, symmetry and Banzhaf value axiomatizations. Int. J. Game Theory 31, 47-67.

McQuillin B (2009) The extended and generalized Shapley value: Simultaneous consideration of coalitional externalities and coalitional structure. J. Econ. Theory 144, 696-721.

McQuillin B, Sugden R (2016) Backward induction foundations of the Shapley value. Econometrica 84, 2265-2280.

Myerson RB (1977) Graphs and cooperation in games. Math. Oper. Res. 2, 225-229.

Myerson RB (1980) Conference structures and fair allocation rules. Int. J. Game Theory 9, 169-182.

Nowak AS (1997) On an axiomatization of the Banzhaf value without the additivity axiom. Int. J. Game Theory 26, 137-141.

Pérez-Castrillo D, Wettstein D (2001) Bidding for the surplus: A non-cooperative approach to the Shapley value. J. Econ. Theory 100, 274–294.

Pham Do KH, Norde H (2007) The Shapley value for partition function form games. Int. Game Theory Rev. 9, 353-360.

Rawls J (1971) A Theory of Justice, Oxford University Press.

Sánchez SF (1997) Balanced contributions axiom in the solution of cooperative games. Games Econ. Behav. 20, 161-168.

Shapley LS (1953) A value for n-person games, in: Kuhn, HW, Tucker, AW (Eds.), Contributions to the Theory of Games, II. Princeton University Press, Princeton, pp. 307-317.

Shapley LS, Shubik, M (1954) A method for evaluating the distribution of power in a committee system. Amer. Polit. Sci. Rev. 48, 787-792.

Shapley LS, Shubik, M (1954) The assignment game I: The core. Int. J. Game Theory 1, 111-130.

Shapley LS, Shubik, M (1969) Pure competition, coalitional power, and fair division. Int. Econ. Rev. 10, 337-362.

Schmeidler D (1969) The nucleolus of a characteristic function game. SIAM J. App. Math. 17, 1163-1170.

Thrall RM, Lucas WF (1963) n-person games in partition function form. Nav. Logist. Q. 10, 281-293.

Wicksell K (1896/1958) A new principle of just taxation. English translation (by JM Buchanan) of German original, in Musgrave, RA, Peacock, AT (eds), Classics in the Theory of Public Finance. Macmillan, London, pp. 72–118.