

# Balanced contributions, consistency, and value for games with externalities

André Casajus<sup>a</sup>, Yukihiko Funaki<sup>b</sup>, Frank Huettner<sup>c,\*</sup>

<sup>a</sup>HHL Leipzig Graduate School of Management, Jahnallee 59, 04109 Leipzig, Germany

<sup>b</sup>School of Political Science and Economics, Waseda University, Nishi-Waseda, Shinjuku-Ku, Tokyo, Japan

<sup>c</sup>SKK GSB, Sungkyunkwan University, 25-2, Sungkyunkwan-Ro, Jongno-gu, 110-745 Seoul, South Korea

## Abstract

We consider fair and consistent extensions of the Shapley value for games with externalities. Based on the restriction identified by Casajus et al. (2024, Games Econ. Behavior 147, 88-146), we define balanced contributions, Sobolev's consistency, and Hart and Mas-Colell's consistency for games with externalities, and we show that these properties lead to characterizations of the generalization of the Shapley value introduced by Macho-Stadler et al. (2007, J. Econ. Theory 135, 339-356), that parallel important characterizations of the Shapley value.

*Keywords:* Shapley value, partition function form, random partition, restriction operator

*2010 MSC:* 91A12, *JEL:* C71, D60

## 1. Introduction

The question of how to divide jointly created value in cooperative games has been fundamentally influenced by the concept of the Shapley value. This solution concept was originally designed to capture a player's value in a cooperative game with transferable utility (henceforth TU game), but its application nowadays transitions into seemingly unrelated domains such as statistics for identifying important variables (Lipovetsky and Conklin, 2001; Shorrocks, 2012) or into machine learning for interpreting prediction models (Lundberg and Lee, 2017; Lundberg et al., 2020). The Shapley value applies under the presupposition that the worth of a coalition is independent of the coalition structure of the outside players. However, this assumption is challenged when external effects come into play, that is, when the actions of one coalition exert influence over another's outcomes. To capture externalities, Thrall and Lucas (1963) introduced the framework of *games with externalities* (henceforth

\*Corresponding author

Email addresses: mail@casajus.de (André Casajus), funaki@waseda.jp (Yukihiko Funaki), mail@frankhuettner.de (Frank Huettner)

URL: www.casajus.de (André Casajus), yfunaki.blogspot.com (Yukihiko Funaki), www.frankhuettner.de (Frank Huettner)

*TUX games*). Crucially, in a TUX game, the worth of a coalition may depend on the coalitions formed by the outsiders, which makes it a generalization of TU games.

TUX games are often used—but are not limited—to study coalition formation (Grabisch and Funaki, 2012; Basso et al., 2021). Indeed, recently Saavedra-Nieves and Casas-Méndez (2023) capture covert networks with TUX games to evaluate agents’ importance in such networks, demonstrating that TUX games can lead to more powerful models and insights than TU games. Whereas the Shapley value probably is the most-prominent single-valued solution concept for TU games, there is a lack of clarity when it comes to solutions for TUX games. Indeed, many *generalizations of the Shapley value* to TUX games were proposed in the literature, usually motivated by characterizations in the spirit of Shapley’s original characterization based on additivity.<sup>1</sup> However, the Shapley value (and derivatives thereof) are often justified by axioms that involve varying player sets, such as the balanced contributions property (Myerson, 1980) and consistency properties deriving from Hart and Mas-Colell (1989) and from Sobolev (1975). Therefore, Dutta et al. (2010) and Casajus et al. (2024) argue that generalizations of the Shapley value should also inherit its properties that relate to varying player sets. After all, these properties not only lie at the heart of many fairness properties of the Shapley value, but also build the fundament for implementations via non-cooperative games (see, e.g., Gul (1989), Macho-Stadler et al. (2007), McQuillin and Sugden (2016), Brügemann et al. (2018)).

### 1.1. Novel axioms for games with externalities and characterizations of the MPW Solution

In this paper, we define a notion of *balanced contributions for games with externalities* along the lines of Myerson (1980), which requires that the impact of removing player  $j$  on player  $i$ ’s payoff is the same as the impact of removing player  $i$  on player  $j$ ’s payoff. We demonstrate that this property together with efficiency is characteristic of the solution introduced by Feldman (1996) and characterized by Macho-Stadler, Pérez-Castrillo and Wettstein (2007), henceforth abbreviated as *MPW solution*, a generalization of the Shapley value (Theorem 4).

Moreover, we investigate the consistency properties introduced by Hart and Mas-Colell (1989) and by Sobolev (1975), which are well-known to be characteristic of the Shapley value together with efficiency and standardness for two-player games. We *augment these consistency properties to TUX games*. Our second main contribution are two novel characterizations of the MPW solution (Theorems 8 and 9). Generalizing Sobolev (1975) to games with externalities, we show that the generalized version of Sobolev’s consistency together with efficiency and standardness for two-player games is characteristic of the MPW solution. Generalizing Hart and Mas-Colell (1989) to games with externalities, we show that the generalized version of Hart and Mas-Colell’s consistency together with standardness for two-player games is characteristic of the MPW solution. In this sense, we contribute to the stream of literature that generalizes these consistency properties to broader frameworks

---

<sup>1</sup>Myerson (1977), Bolger (1989), Albizuri et al. (2005), Pham Do and Norde (2007), Macho-Stadler et al. (2007), McQuillin (2009), Dutta et al. (2010), Grabisch and Funaki (2012), and Skibski et al. (2018) introduce different (classes of) solutions.

with the intend to generalize the Shapley value (see Winter (1992), Dutta et al. (2010), or Xu et al. (2013) for other such attempts).

## 1.2. Relation to the Literature

Several motivations for studying balanced contributions and consistency arise from the literature. We will give a brief summary.

### 1.2.1. *Balanced contributions property is a fairness property that is fruitful for derived axiomatizations*

Unlike additivity, the balanced contributions property is a fairness statement (Moulin, 2003). It establishes an equitable relationship between players' contributions to each other. This property has proven instrumental in characterizing various extensions of the Shapley value (e.g., Lorenzo-Freire et al. (2007), Gómez-Rúa and Vidal-Puga (2010), Kamijo and Kongo (2012), and van den Brink et al. (2014)). The balanced contributions property was also key for motivating specific applications of the Shapley value, including less traditional applications in machine learning (Davila-Pena et al., 2022), as well as applications in more traditional realms. Recent examples involve network games (González Arangüena et al., 2015), revenue sharing (Bergantiños and Moreno-Ternero, 2025), and group decision models (Meng et al., 2023). These applications demonstrate how balanced contributions can effectively embody core fairness principles when adapted to particular operational contexts. It further suggests that extending this axiom to games with externalities can lead to a plethora of derivative insights.

### 1.2.2. *Merits of consistency properties*

Exploring the consistency properties within cooperative game theory serves a tripartite purpose (Driessen, 1991). First, it enables a clear differentiation between various solution concepts, offering a nuanced perspective on their unique properties and conditions of application. For comparison, Álvarez-Mozos and Ehlers (2024) provide a characterization of the prenucleolus for TUX games based on a consistency property generalizing the notion of the reduced game introduced by Davis and Maschler (1965). Second, exploring consistency properties fosters the theoretical evolution of these solution concepts, deepening our understanding of their underpinnings and potential for refinement. This is, for example, important for the development of mechanisms that implement a solution concept by a non-cooperative game. To this end, it is particularly useful to know about ways to pay out players and reduce the size or complexity of a game without changing the payoffs of the remaining players. Prominent examples for mechanisms that implement the Shapley value and that stepwise reduce the number of players are given by Macho-Stadler et al. (2007), McQuillin and Sugden (2016), and McQuillin and Sugden (2018); implementations of derivatives of the Shapley value along such lines are given by, for example, Bergantiños and Vidal-Puga (2010), Ju et al. (2014), and Bergantiños et al. (2023). Lastly, the consistency property may be useful in order to determine coherent solutions for realistic problems, that is, when the game derives from specific applications. Important examples for which consistency properties of TU games boil down to meaningful properties for the specific application include

airport problems (Littlechild and Owen, 1973; Potters and Sudhölter, 1999); bankruptcy problems (O’Neill, 1982; Aumann and Maschler, 1985; Thomson, 2015); sequencing problems (Curiel et al., 1989; van den Brink and Chun, 2011); highway problems (Kuipers et al., 2013; Sudhölter and Zarzuelo, 2017); queuing problems (Maniquet, 2003; Bendel and Haviv, 2018; Thomson and Velez, 2022); minimal cost spanning tree problems (Dutta and Kar, 2004); probabilistic assignment (Chambers, 2004); and allocating greenhouse gas emission costs (Gopalakrishnan et al., 2021). Even though such allocation problems mostly lend themselves to be studied with externalities, this has largely been committed from the analysis so far. Our work introduces a tool that incorporates both externalities and consistency principles for a general model of allocation problems, which allows for a reexamination of these models that includes externalities.

### *1.2.3. Characterizations of the MPW solution*

The MPW solution was first introduced by Feldman (1996) and characterized by Macho-Stadler et al. (2007) who use Shapley’s classical axioms of linearity, efficiency, and the dummy player property in combination with a strengthening of the symmetry property and a similar influence requirement. Fujinaka (2004) provides characterizations of solutions—among others of the MPW solution—in the spirit of Young (1985), ensuring that each player’s reward depends only on this player’s vector of average marginal contributions. Skibski et al. (2018) highlight the relationship of the Ewens distribution to the “Chinese restaurant process” (Alldous, 1985, 11.19; Pitman, 2006, Equation 3.3) and that the MPW solution emerges as the expected marginal contribution of this stochastic coalition formation process. Skibski and Michalak (2019) deduce a characterization based on Shapley’s classical axioms and on properties of the probability distributions employed in the stochastic coalition formation process.

Casajus et al. (2024) argue that the MPW solution is the only plausible generalization of the Shapley value that—like the Shapley value—admits a potential (Hart and Mas-Colell, 1989), that in turn can be obtained as an expected accumulated worth of partitions (Casajus, 2014). We contribute to this literature by providing novel characterizations based on equal-gains and consistency principles.

### *1.3. Structure the paper*

The remainder of this paper is structured as follows. In Section 2, we introduce basic definitions for TU games and TUX games. In Section 3, we revisit the balanced contributions property characteristic of the Shapley value, augment balanced contributions to TUX games, and use it for a characterization of the MPW solution. In Section 4, we revisit the consistency properties characteristic of the Shapley value, generalize these consistency properties to TUX games, and use them for novel characterizations of the MPW solution. We end with some concluding remarks. The appendix contains all the proofs.

## 2. Basic definitions and notation

Let  $\mathbf{U}$  be a finite set of players, the **universe of players**. For  $N \subseteq \mathbf{U}$ , let  $2^N$  denote the set of all subsets of  $N$ .<sup>2</sup> Throughout the paper, the cardinalities of coalitions  $N, S, T, B \subseteq \mathbf{U}$  are denoted by  $n, s, t$ , and  $b$ , respectively.

### 2.1. Games without externalities and the Shapley value

A cooperative game with transferable utility, henceforth **TU game** (also known as a **game in characteristic function form**), for a player set  $N \subseteq \mathbf{U}$  is given by its **characteristic function**  $v : 2^N \rightarrow \mathbb{R}$ ,  $v(\emptyset) = 0$ , which assigns a worth to each coalition  $S \subseteq N$ . Let  $\mathbb{V}(N)$  denote the set of all TU games for  $N$  and let  $\mathbb{V}$  denote the set of all TU games. For  $v \in \mathbb{V}(N)$ ,  $N \subseteq \mathbf{U}$ , and  $T \subseteq N$ , the **restriction** of  $v$  by **removing the players** in  $T$ ,  $v_{-T} \in \mathbb{V}(N \setminus T)$ , is given by  $v_{-T}(S) = v(S)$  for all  $S \subseteq N \setminus T$ . Alternatively, one can address the game  $v_{-T}$  as the **subgame** of  $v$  on the player set  $N \setminus T$ .

A **solution for TU games** is an operator  $\varphi$  that assigns a payoff vector  $\varphi(v) \in \mathbb{R}^N$  to any TU game  $v \in \mathbb{V}(N)$ ,  $N \subseteq \mathbf{U}$ . The **Shapley value** (Shapley, 1953),  $\text{Sh}$ , is given by

$$\text{Sh}_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{s!(n-s-1)!}{n!} (v(S \cup \{i\}) - v(S)) \quad (1)$$

for all  $v \in \mathbb{V}(N)$ ,  $N \subseteq \mathbf{U}$  and  $i \in N$ .

### 2.2. Games with externalities and the MPW solution

A **partition** of  $N \subseteq \mathbf{U}$  is a collection of non-empty subsets of  $N$  such that any two of them are disjoint and such their union is  $N$ . The set of partitions of  $N \subseteq \mathbf{U}$  is denoted by  $\Pi(N)$ . For technical reasons, we set  $\Pi(\emptyset) = \{\emptyset\}$ . The block of  $\pi \in \Pi(N)$  that contains player  $i \in N$  is denoted by  $\pi(i)$ .

For  $\pi \in \Pi(N)$ ,  $N \subseteq \mathbf{U}$ , the **elimination of the players** in  $T \subseteq N$  from  $\pi$  gives  $\pi_{-T} \in \Pi(N \setminus T)$ ,

$$\pi_{-T} = \{\{B \setminus T) \mid B \in \pi \text{ and } B \setminus T \neq \emptyset\}.$$

Instead of  $\pi_{-\{i\}}$ , we write  $\pi_{-i}$ . For  $N \subseteq \mathbf{U}$  and  $\pi \in \Pi(N)$ , **adding a player**  $i \in \mathbf{U} \setminus N$  to the block  $B \in \pi$  is denoted by  $\pi_{+i \sim B} \in \Pi(N \cup \{i\})$ ,

$$\pi_{+i \sim B} = (\pi \setminus \{B\}) \cup \{B \cup \{i\}\};$$

adding player  $i$  as a singleton is denoted by  $\pi_{+i \sim \emptyset} \in \Pi(N \cup \{i\})$ ,  $\pi_{+i \sim \emptyset} = \pi \cup \{\{i\}\}$ .

A TU game with externalities, henceforth **TUX game** (also known as **game in partition function form**), for a player set  $N \subseteq \mathbf{U}$  is given by its **partition function**  $w : \mathcal{E}(N) \rightarrow \mathbb{R}^N$ , where  $\mathcal{E}(N)$  denotes the set of **embedded coalitions**  $(S, \pi)$  for  $N$  given by

$$\mathcal{E}(N) = \{(S, \pi) \mid S \subseteq N \text{ and } \pi \in \Pi(N \setminus S)\},$$

---

<sup>2</sup>Note that we do not add players in the proofs, so that a finite universe of players is sufficient.

and with  $w(\emptyset, \pi) = 0$  for all  $\pi \in \Pi(N)$ . We denote the set of all TUX games for a player set  $N$  by  $\mathbb{W}(N)$  and the set of all TUX games by  $\mathbb{W}$ . For  $N \subseteq \mathbf{U}$ , the **null game**  $\mathbf{0}^N \in \mathbb{W}(N)$  is defined by  $\mathbf{0}^N(S, \pi) = 0$  for all  $(S, \pi) \in \mathcal{E}(N)$ .

### 2.2.1. The MPW solution

Several solutions were introduced in the literature to generalize the Shapley value, i.e., solutions that boil down to the Shapley value if a TUX game actually does not exhibit externalities, i.e., if  $w(S, \pi) = w(S, \tau)$  for all  $S \subseteq N$  and  $\pi, \tau \in \Pi(N \setminus S)$ . Macho-Stadler et al. (2007) put forth a solution for TUX games, MPW, following a two step procedure:<sup>3</sup>

1. For a given TUX game  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$ , one first computes a TU game  $\bar{v}_w \in \mathbb{V}(N)$ , the average game, in which each coalition  $S$  gets the following expected value of  $w(S, \pi)$  over all partitions  $\pi \in \Pi(N \setminus S)$ ,

$$\bar{v}_w(S) = \sum_{\pi \in \Pi(N \setminus S)} \frac{\prod_{B \in \pi} (b-1)!}{(n-s)!} w(S, \pi) \quad \text{for all } S \subseteq N. \quad (2)$$

2. Second, one applies the Shapley value to this TU game in order to obtain the MPW solution,

$$\text{MPW}(w) = \text{Sh}(\bar{v}_w) \quad \text{for all } w \in \mathbb{W}(N), N \subseteq \mathbf{U}. \quad (3)$$

Note that

$$p_{N \setminus S}^*(\pi) = \prod_{B \in \pi} \frac{(b-1)!}{(n-s)!} \quad (4)$$

for all  $S \subseteq N \subseteq \mathbf{U}$  and  $\pi \in \Pi(N \setminus S)$  is a probability distribution over the set of partitions  $\Pi(N \setminus S)$ . A family of probability distributions  $p = (p_N)_{N \subseteq \mathbf{U}}$  over partitions is called a random partition for  $\mathbf{U}$ . The random partition  $p^*$  is known as the Ewens distribution with mutation rate  $\theta = 1$  (Ewens, 1972), which takes a central role in the literature on random partitions (Crane, 2016).

### 2.2.2. Subgames for TUX games

For TU games, there is an obvious way to obtain subgames. In contrast, the notion of a subgame is less obvious for TUX games, since we cannot simply read it off the original game. When player  $i$  is removed from the TUX game  $w$ , we have to specify the worth of each embedded coalition “ $w_{-i}(S, \pi)$ ” in the TUX game  $w_{-i}$  without player  $i$ . For instance, when removing player 4 from some TUX game  $w \in \mathbb{W}(\{1, 2, 3, 4\})$ , the worth  $w_{-4}(\{1\}, \{2, 3\})$  has no obvious reference in the original game  $w$ , where player 4 impacts the worth of coalition  $\{1\}$  through being singleton or being affiliated with  $\{2, 3\}$ .

To capture the many possibilities of how to obtain subgames in the presence of externalities, Dutta et al. (2010) introduce the concept of a **restriction operator**. A restriction

---

<sup>3</sup>The MPW solution is part of a larger class called average Shapley values introduced and characterized by Macho-Stadler et al. (2007), which includes solutions derived by an analogous two step procedure but with possibly other probability distributions over  $\Pi(N \setminus S)$ .

operator  $r$  formally specifies how to obtain a “subgame”  $w_{-i}^r \in \mathbb{W}(N \setminus \{i\})$  for every TUX game  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$ , and every player  $i \in N$ , where the worth  $w_{-i}^r(S, \pi)$  in the subgame only depends on the worths  $w(S, \pi_{+i \rightsquigarrow \emptyset})$  and  $w(S, \pi_{+i \rightsquigarrow B})$  for  $B \in \pi$  in the original game  $w$ . For example, a restriction operator  $r$  specifies how the worth  $w_{-4}^r(\{1\}, \{\{2, 3\}\})$  is derived from aggregating the numbers  $w(\{1\}, \{\{2, 3\}, \{4\}\})$  and  $w(\{1\}, \{\{2, 3, 4\}\})$ . Subgames that are constructed in this way can be seen as “‘estimates’ or ‘approximations’ based on the available data” (Dutta et al., 2010).

To ensure that subgames are well-defined even if more than one player was removed, i.e., to consider  $w_{-T}^r$ ,  $T \subseteq N$ , it must not matter in which order the players are removed,  $(w_{-i}^r)_{-j}^r = (w_{-j}^r)_{-i}^r$  for all  $N \subseteq \mathbf{U}$ ,  $w \in \mathbb{W}(N)$ , and  $i, j \in N$ ,  $i \neq j$ . A restriction operator satisfying this principle is said to be path independent.

Whilst there exists a plethora of possible path independent restrictions operators, Casajus et al. (2024) argue there is a unique way to define a restriction operator that maintains various properties of the Shapley value.

### 2.2.3. The restriction operator $r^*$

Hart and Mas-Colell (1989) show that the Shapley value is a player’s contribution to the so-called potential of a game, which can be obtained as the expected accumulated worth under a probabilistic partitioning scheme (Casajus, 2014). Specifically, given a suitable probability distribution defined over the set of all possible partitions of the player set, the potential function equals the expected value of the sum of coalition worths within each randomly selected partition. Imposing that a potential for TUX games maintains this interpretation, and resting on the fact every path-independent restriction operator corresponds to a potential for TUX games (Dutta et al., 2010, Theorem 1), Casajus et al. (2024) single out a unique restriction operator  $r^*$ .<sup>4</sup>

In what follows, we will limit the analysis to the restriction operator  $r^*$ , and we will omit explicit mention of the superscript. It is defined by

$$w_{-i}(S, \pi) = w_{-i}^{r^*}(S, \pi) = \frac{1}{n-s} \left( w(S, \pi_{+i \rightsquigarrow \emptyset}) + \sum_{j \in N \setminus (S \cup \{i\})} w(S, \pi_{+i \rightsquigarrow \pi(j)}) \right) \quad (5)$$

for all  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$ ,  $i \in N$ , and  $(S, \pi) \in \mathcal{E}(N \setminus \{i\})$ . Thus, to compute worth of an embedded coalition  $(S, \pi)$  without player  $i$ , we perform a counterfactual analysis. This analysis considers what would happen if player  $i$  were not explicitly part of coalition  $S$ . To do this, we imagine player  $i$  engaging in all possible interactions with players outside of  $S$ , or choosing to remain isolated. Giving equal chance to player  $i$  making connections with any individual player  $j \in N \setminus S$  who is not in coalition  $S$  (making a connection with player  $i$  themselves is interpreted as staying alone), we expect the worth  $w_{-i}(S, \pi)$ . This simple average over the worths of embedded coalitions then reflects the potential externalities due

---

<sup>4</sup>Casajus et al. (2024) impose further assumptions to ensure that the obtained potential for TUX games generalizes the potential for TU games.

to player  $i$ . For example, removing player 4 from  $w \in \mathbb{W}(\{1, 2, 3, 4\})$  gives

$$w_{-4}^r(\{1\}, \{\{2, 3\}\}) = \frac{1}{3}w(\{1\}, \{\{2, 3\}, \{4\}\}) + \frac{2}{3}w(\{1\}, \{\{2, 3, 4\}\}).$$

It is important to note that a removed player casts a shadow on the subgame. For example, a subgame originating from a four-player Cournot oligopoly is distinct from a three-firm Cournot oligopoly.

**Example 1.** Consider a Cournot oligopoly with four firms  $N = \{1, 2, 3, 4\}$ , which have identical constant marginal cost  $c > 0$ , and which face inverse demand  $P(X) = A - X_N$  with  $A > c$ , where the output of some coalition  $S \subseteq N$  is given by  $X_S = \sum_{i \in S} x_i$ . Given a partition  $\pi \in \Pi(N)$  into  $|\pi|$  cartels, each cartel  $S \in \pi$  chooses the joint quantity  $X_S$  to maximizes its joint profit  $(A - c - X_{N \setminus S} - X_S)X_S$ . In the Cournot equilibrium, each cartel  $S \in \pi$  then has the same profit

$$w(S, \pi \setminus \{S\}) = \frac{(A - c)^2}{(|\pi| + 1)^2}.$$

Note that profits are larger the less cartels there are, i.e., merging cartels exercise positive external effects on the other cartels (see Yi (1997) for a more precise definition of positive/negative external effects in this context).

Reducing this TUX game to a subgame now means capturing the strategic interactions by a TUX game with the remaining firms. The possibility of external effects of the removed firm is still accounted for, but the removed firm's ability to actively participate in coalition formation with the remaining firms is no longer considered. Instead, when modeling the profit of a cartel in the subgame, we take a probabilistic approach concerning the alliance of the removed firm with outside firms. The restriction without firm 4,  $w_{-4} \in \mathbb{W}(\{1, 2, 3\})$  is given by

$$\begin{aligned} w_{-4}(\{i\}, \{\{j\}, \{k\}\}) &= \frac{1}{3}w(\{i\}, \{\{j\}, \{k\}, \{4\}\}) + \frac{2}{3}w(\{i\}, \{\{j, 4\}, \{k\}\}) = \frac{11}{200}(A - c)^2 \\ w_{-4}(\{i\}, \{\{j, k\}\}) &= \frac{1}{3}w(\{i\}, \{\{j, k\}, \{4\}\}) + \frac{2}{3}w(\{i\}, \{\{j, k, 4\}\}) = \frac{41}{432}(A - c)^2 \\ w_{-4}(\{i, j\}, \{\{k\}\}) &= \frac{1}{2}w(\{i, j\}, \{\{k\}, \{4\}\}) + \frac{1}{2}w(\{i, j\}, \{\{k, 4\}\}) = \frac{25}{288}(A - c)^2 \\ w_{-4}(\{1, 2, 3\}, \{\emptyset\}) &= w(\{1, 2, 3\}, \{\{4\}\}) = \frac{1}{9}(A - c)^2 \end{aligned}$$

for  $\{i, j, k\} = \{1, 2, 3\}$ . We observe that  $w_{-4}(\{i\}, \{\{j, k\}\}) > w_{-4}(\{i, j\}, \{\{k\}\})$ . This difference arises from the higher probability that firm 4 is believed to connect with  $\{j, k\}$  compared to  $\{k\}$ . Consequently, the likelihood of firm 4 operating independently (and thus not contributing positive externalities to the embedded coalition) is reduced. Clearly, this

subgame differs from a standard three-firm Cournot oligopoly,  $N = \{1, 2, 3\}$ . For instance,

$$w(\{1, 2, 3\}) = \frac{1}{4}(A - c)^2 > \frac{1}{9}(A - c)^2 = w_{-4}(\{1, 2, 3\}, \{\emptyset\}).$$

The subgame  $w_{-4}$  acknowledges the negative externalities exerted by firm 4 on the profit of the grand coalition  $\{1, 2, 3\}$ . Therefore, while this restriction accounts for the removed firm's externalities, it simultaneously constrains any potential cartel formation with that firm.

### 3. Balanced contributions

We start by revisiting the definition of balanced contributions for TU games and the characterization of the Shapley value based on this property. Thereafter, we extend the result to TUX games.

#### 3.1. Myerson's characterization of the Shapley value

Myerson (1980) observed that the Shapley value satisfies the following property.

*Balanced Contributions, BC.* For all  $v \in \mathbb{V}(N)$ ,  $N \subseteq \mathbf{U}$ , and  $i, j \in N$ , we have  $\varphi_i(v) - \varphi_i(v_{-j}) = \varphi_j(v) - \varphi_j(v_{-i})$ .

According to this requirement the impact of removing player  $j$  on player  $i$ 's payoff is the same as the impact of removing player  $i$  on player  $j$ 's payoff. There is a rich literature investigating the consequences of this property. Equivalent properties are summarized by Casajus and Huettner (2018). Not much is missing to obtain a characterization of the Shapley value.

*Efficiency, EF.* For all  $v \in \mathbb{V}(N)$ ,  $N \subseteq \mathbf{U}$ , we have  $\sum_{i \in N} \varphi_i(v) = v(N)$ .

Perhaps surprisingly, balanced contributions together with efficiency already is characteristic of the Shapley value.

**Theorem 2 (Myerson, 1980).** *The Shapley value,  $\text{Sh}$ , is the unique solution for TU games that satisfies efficiency (EF) and the balanced contributions property (BC).*

Next, we turn to the case with externalities.

#### 3.2. A characterization of the MPW solution based on balanced contributions

Based on the restriction defined in (5), we obtain a straight-forward generalization of the balanced contributions property to TUX games.

*Balanced Contributions, BC<sup>X</sup>.* For all  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$ , and  $i, j \in N$ , we have  $\varphi_i(w) - \varphi_i(w_{-j}) = \varphi_j(w) - \varphi_j(w_{-i})$ .

The impact of removing player  $j$  on player  $i$ 's payoff is the same as the impact of removing player  $i$  on player  $j$ 's payoff, when removing a player is accounted for in the fashion proposed by Casajus et al. (2024). Consider the following example introduced by Maskin (2003) to get a better intuition for this property.

**Example 3.** Let a three-player public goods game be given by  $N = \{1, 2, 3\}$  and

$$\begin{aligned} w(\{i\}, \{\{j\}, \{k\}\}) &= 0; \\ w(\{i\}, \{\{j, k\}\}) &= 9; \\ w(\{1, 2\}, \{\{3\}\}) &= 12; \\ w(\{1, 3\}, \{\{2\}\}) &= 13; \\ w(\{2, 3\}, \{\{1\}\}) &= 14; \\ w(\{1, 2, 3\}, \{\emptyset\}) &= 24. \end{aligned}$$

The subgames  $w_{-i}$ ,  $i \in \{1, 2, 3\}$  reflect that an affiliation of the removed player with outside coalitions induces externalities,

$$w_{-i}(\{j\}, \{\{k\}\}) = \frac{1}{2}w(\{i\}, \{\{j\}, \{k\}\}) + \frac{1}{2}w(\{i\}, \{\{j, k\}\}) = 4.5.$$

The subgames also reflect the asymmetries between the players:  $w_{-1}(\{23\}, \{\emptyset\}) = 14$ ,  $w_{-2}(\{13\}, \{\emptyset\}) = 13$ , and  $w_{-3}(\{12\}, \{\emptyset\}) = 12$ .

TUX games with two players are in fact TU games. Applying the Shapley value to the restricted games gives the payoffs  $\varphi_1(w_{-2}) = 6.5$ ,  $\varphi_1(w_{-3}) = 6$ ,  $\varphi_2(w_{-1}) = 7$ , and  $\varphi_3(w_{-1}) = 7$ . Balanced contributions addresses these asymmetries and requires equal gains,

$$\begin{aligned} \varphi_1(w) - 6.5 &= \varphi_2(w) - 7; \\ \varphi_1(w) - 6 &= \varphi_3(w) - 7. \end{aligned}$$

Note that together with efficiency,  $\varphi_1(w) + \varphi_2(w) + \varphi_3(w) = 24$ , we have enough independent equations to infer  $\varphi_1(w) = 7.5$ ,  $\varphi_2(w) = 8$ , and  $\varphi_3(w) = 8.5$ .

In this example, the payoffs coincide with the efficient generalized Shapley value introduced by Hafalir (2007). In contrast, the payoffs according to the “externality-free value” (de Clippel and Serrano, 2008), which was introduced by Pham Do and Norde (2007), follow from the assumption that outside players always stay singletons (hence, it ignores the data  $w(\{i\}, \{\{j, k\}\}) = 9$ ) and are given by  $(\frac{43}{6}, \frac{44}{6}, \frac{45}{6})$ .

It turns out that the balanced contributions property together with efficiency is characteristic of the MPW solution. We prepare the statement of this result with a formal definition of efficiency for TUX games.

**Efficiency,  $\mathbf{EF}^X$ .** For all  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$ , we have  $\sum_{i \in N} \varphi_i(w) = w(N, \emptyset)$ .

We can now state our first main result.

**Theorem 4.** *The MPW solution, MPW, is the unique solution for TUX games that satisfies efficiency ( $\mathbf{EF}^X$ ) and the balanced contributions property ( $\mathbf{BC}^X$ ).*

Whereas the proof of uniqueness rests on an inductive argument, it may be less obvious that the MPW solution actually satisfies the balanced contributions property. However, this is a consequence of the fact that the Shapley value satisfies the balanced contributions

property, the fact that the MPW solution is the Shapley value of the average game, and the following lemma, which states that it does not matter if we first average a game and then remove a player from the obtained TU game,  $(\bar{v}_w)_{-i}$ ; or whether we first remove a player from the TUX game and then take the average second  $\bar{v}_{(w-i)}$ .

**Lemma 5.** *For all  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$ , and  $i \in N$ , we have  $(\bar{v}_w)_{-i} = \bar{v}_{(w-i)} \in \mathbb{V}(N \setminus \{i\})$ .*

It turns out that this commutative relationship of removal and average operators is instrumental for further results as demonstrated in the next section.

#### 4. Consistency

An important stream of characterization results of the Shapley value and alternative solution concepts for TU games relies on consistency properties. These properties involve the notion of a reduced game, which is formed when one or more players are removed from the original game, with the understanding that these players receive compensation as determined by a specific payoff principle. The reduced game thus depends on the original game, the payoffs allocated to the removed players, and (sometimes) it also depends on the solution proposed for a subgame. The consistency property asserts that if a payoff vector exists for the original game, then a corresponding payoff vector should be achievable for the reduced game's players—ensuring a seamless transition in value distribution despite changes in player configuration.

For the characterization of the Shapley value, Peleg and Sudhölter (2007) highlight the consistency properties introduced by Sobolev (1975) and Hart and Mas-Colell (1989), which we revisit next. Thereafter, we introduce consistency properties that generalize these properties to TUX games.

##### 4.1. Sobolev's and Hart and Mas-Colell's characterization of the Shapley value

Among the consistency properties for TU games (without externalities) discussed in the literature, the notion introduced by Sobolev (1975) has the particular appeal of *not* referring to the payoffs of subgames; instead, it refers to some sort of a protocol. In this sense, it appears to capture the nature of the Shapley value particularly well.

Let  $\varphi$  be a solution for TU games, and let  $v \in \mathbb{V}(N)$ ,  $N \subseteq \mathbf{U}$ , and  $j \in N$ . Sobolev's notion of a reduced TU game  $v_{-j}^{\varphi} \in \mathbb{V}(N \setminus \{j\})$  is given by

$$v_{-j}^{So,\varphi}(S) = \frac{s}{n-1} (v(S \cup \{j\}) - \varphi_j(v)) + \left(1 - \frac{s}{n-1}\right) v(S) \quad \text{for all } S \subseteq N \setminus \{j\}. \quad (6)$$

The idea behind this reduced game can be described as follows. Player  $j$  will join forces with one of the  $n-1$  other players, each player equally likely. The worth of a coalition  $S$  in the reduced game depends on whether player  $j$  joins one of the  $s$  players within coalition  $S$ . If this is the case, then coalition  $S$  makes use of player  $j$ 's productivity and compensates player  $j$  with  $\varphi_j(v)$ , so that the worth of coalition  $S$  equals  $v(S \cup \{j\}) - \varphi_j(v)$ ; otherwise, player  $j$  joins the other players and the worth of coalition  $S$  remains the same. In expectation,

coalition  $S$  has the worth  $v_{-j}^{So,\varphi}(S)$ . Interestingly, such a reduction or amalgamation of player  $j$  is neutral to the Shapley payoff of the other players, that is, the Shapley value satisfies the following property.

*Sobolev Consistency, SC.* For all  $v \in \mathbb{V}(N)$ ,  $N \subseteq \mathbf{U}$ , and  $i, j \in N$ ,  $i \neq j$ , we have  $\varphi_i(v) = \varphi_i(v_{-j}^{So,\varphi})$ .

It is further useful to compare this with the consistency property proposed by Hart and Mas-Colell (1989). To this end, define the reduced game  $v_{-j}^{HM,\varphi} \in \mathbb{V}(N \setminus \{j\})$  by

$$v_{-j}^{HM,\varphi}(S) = v(S \cup \{j\}) - \varphi_j(v_{-N \setminus (S \cup \{j\})}) \quad \text{for all } S \subseteq N \setminus \{j\}.$$

Here, the worth of coalition  $S$  in a subgame is obtained by utilizing the productivity of the removed player  $j$  and paying to  $j$  the payoff that player  $j$  obtains in the game without the other players  $v_{-N \setminus (S \cup \{j\})}$ . Again, the Shapley value applied to the reduced game gives a player the same payoff as in the original game, that is, the Shapley value satisfies the following property.

*Hart and Mas-Colell Consistency, HMC.* For all  $v \in \mathbb{V}(N)$ ,  $N \subseteq \mathbf{U}$ , and  $i, j \in N$ ,  $i \neq j$ , we have  $\varphi_i(v) = \varphi_i(v_{-j}^{HM,\varphi})$ .

Different from the Sobolev consistency, HM consistency applies the solution concept to a subgame  $v_{-S \setminus T}$  in the definition of the reduced game.

The Shapley value is the only efficient solution concept that satisfies these consistency properties and standardness for two-player games, which simply prescribes that the surplus in a two-player game is equally shared.

*2-Standardness, 2S.* For all  $v \in \mathbb{V}(\{i, j\})$ ,  $\{i, j\} \subseteq \mathbf{U}$ ,  $i \neq j$ , we have

$$\varphi_i(v) = v(\{i\}) + \frac{v(\{i, j\}) - v(\{j\}) - v(\{i\})}{2}.$$

The literature emphasizes the following characterizations.

**Theorem 6 (Sobolev, 1975).** *The Shapley value, Sh, is the unique solution that satisfies efficiency (EF), 2-standardness (2S), and Sobolev consistency (SC).*

**Theorem 7 (Hart and Mas-Colell, 1989).** *The Shapley value, Sh, is the unique solution that satisfies 2-standardness (2S) and Hart and Mas-Colell consistency (HMC).*

We further remark that Hart and Mas-Colell (1989) also discuss the following equivalent notion of HM consistency that allows for the removal of multiple players:

For all  $v \in \mathbb{V}(N)$ ,  $N \subseteq \mathbf{U}$ , and  $T \subseteq N$ ,  $i \in N \setminus T$ , we have

$$\varphi_i(v) = \varphi_i(v_{-T}^{HM,\varphi}), \tag{7}$$

where  $v_{-T}^{HM,\varphi}$  is given

$$v_{-T}^{HM,\varphi}(S) = v(S \cup T) - \sum_{j \in T} \varphi_j(v_{-N \setminus (S \cup T)}) \quad \text{for all } S \subseteq N \setminus T. \quad (8)$$

Note that for  $T = \{j\}$ , the reduced game in (8) becomes  $v_{-T}^{HM,\varphi} = v_{-j}^{HM,\varphi}$ , and the above property (7) boils down to HM consistency; hence, (7) implies HM consistency. Perhaps surprisingly, HM consistency in turn implies property (7) even if  $t > 1$  (Hart and Mas-Colell, 1989, Lemma 4.4). This is a key to establish the characterization in Theorem 7.

#### 4.2. Sobolev consistency for TUX Games

When generalizing Sobolev's consistency, we have to be careful with the formulation of the reduced game. Note that in the definition of the reduced game, (6), the worth  $v(S)$  conceptually refers to the worth in the game without player  $j$ , that is, to a subgame. In order to introduce an analogous notion of a reduced game for TUX games, we need to apply the restriction at this place.

Let  $\varphi$  be a solution for TUX games and let  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$ , and  $j \in N$ . The reduced TUX game  $w_{-j}^{So,\varphi} \in \mathbb{W}(N \setminus \{j\})$  is defined by

$$w_{-j}^{So,\varphi}(S, \pi) = \frac{s}{n-1} (w(S \cup \{j\}, \pi) - \varphi_j(w)) + \left(1 - \frac{s}{n-1}\right) w_{-j}(S, \pi) \quad (9)$$

for all  $(S, \pi) \in \mathcal{E}(N \setminus \{j\})$ , where the subgame  $w_{-j}$  is defined in (5). This suggests the following consistency property.

**Sobolev Consistency, SC<sup>X</sup>.** For all  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$ , and  $i, j \in N$ ,  $i \neq j$ , we have  $\varphi_i(w) = \varphi_i(w_{-j}^{So,\varphi})$ .

To illustrate this property, recall the public good game in Example 3, for which  $\text{MPW}_1(w) = 7.5$  and  $w_{-1}(\{2\}, \{\{3\}\}) = 4.5$ . The reduced game without player 1 is given by

$$\begin{aligned} w_{-1}^{So,\varphi}(\{2\}, \{\{3\}\}) &= \frac{1}{2} (w(\{1, 2\}, \{\{3\}\}) - \text{MPW}_1(w)) + \frac{1}{2} w_{-1}(\{2\}, \{\{3\}\}) = 4.5; \\ w_{-1}^{So,\varphi}(\{3\}, \{\{2\}\}) &= 5; \\ w_{-1}^{So,\varphi}(\{23\}, \{\emptyset\}) &= w(\{1, 2, 3\}, \pi) - \text{MPW}_1(w) = 16.5. \end{aligned}$$

Indeed, solving this two-player game gives  $\text{MPW}_2(w_{-1}^{So,\varphi}) = 8$  and  $\text{MPW}_3(w_{-1}^{So,\varphi}) = 8.5$ , which coincides with the payoffs in the original game  $\text{MPW}_2(w)$  and  $\text{MPW}_3(w)$ , respectively.

It is not obvious that the MPW solution satisfies this requirement. We deduce this from the fact that the Shapley value satisfies Sobolev consistency, the fact that the MPW solution is the Shapley value on the average game, and Lemma 5, which states that it does not matter whether we first average a game and then remove a player from the obtained TU game, or whether we first remove a player from the TUX game  $w$  and then take the average, i.e.,  $(\bar{v}_w)_{-i} = \bar{v}_{(w_{-i})}$ .

Whereas the consistency property requires some work to be generalized to TUX games, the generalizations of the other axioms used by Sobolev are straight-forward.

**2-Standardness,  $\mathbf{2S}^{\mathbf{X}}$ .** For all  $w \in \mathbb{W}(\{i, j\})$ ,  $\{i, j\} \subseteq \mathbf{U}$ ,  $i \neq j$ , we have

$$\varphi_i(w) = v(\{i\}, \{\{j\}\}) + \frac{v(\{i, j\}, \emptyset) - v(\{j\}, \{\{i\}\}) - v(\{i\}, \{\{j\}\})}{2}.$$

Analogous to Sobolev (1975), we find that the MPW solution is characterized by these properties.

**Theorem 8.** *The MPW solution, MPW, is the unique solution that satisfies efficiency ( $\mathbf{EF}^{\mathbf{X}}$ ), 2-standardness ( $\mathbf{2S}^{\mathbf{X}}$ ), and Sobolev consistency ( $\mathbf{SC}^{\mathbf{X}}$ ).*

The proof of uniqueness proceeds by induction on the size of the player set (see Appendix A.5).

#### 4.3. Hart and Mas-Colell consistency for TUX games

Next, we consider the notion of HM consistency for games with externalities based on the restriction given in (5). We define the reduced game as follows. Let  $\varphi$  a be solution for TUX games and let  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$ , and  $j \in N$ . The reduced TUX game  $w_{-j}^{HM, \varphi} \in \mathbb{W}(N \setminus \{j\})$  is defined by

$$w_{-j}^{HM, \varphi}(S, \pi) = w(S \cup \{j\}, \pi) - \varphi_j(w_{-N \setminus (S \cup \{j\})}) \quad (10)$$

for all  $(S, \pi) \in \mathcal{E}(N \setminus \{j\})$ , where  $w_{-N \setminus (S \cup \{j\})}$  is defined in (5). This mimics the reduced game for TU games, where a coalition  $S$  can utilize player  $j$  but has to compensate this player according to the solution  $\varphi$  in the subgame with player set  $S \cup \{j\}$ . This reduction suggests the following property.

**HM Consistency,  $\mathbf{HMC}^{\mathbf{X}}$ .** For all  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$ , and  $i, j \in N$ ,  $i \neq j$ , we have  $\varphi_i(w) = \varphi_i(w_{-j}^{HM, \varphi})$ .

To illustrate this property, recall the public good game in Example 3, for which  $\text{MPW}_1(w_{-2}) = 6.5$ ,  $\text{MPW}_1(w_{-3}) = 6$ , and  $\text{MPW}_1(w) = 7.5$ . The reduced game without player 1 is given by

$$\begin{aligned} w_{-1}^{HM, \varphi}(\{2\}, \{\{3\}\}) &= (w(\{2, 1\}, \{\{3\}\}) - \text{MPW}_1(w_{-3})) = 6; \\ w_{-1}^{HM, \varphi}(\{3\}, \{\{2\}\}) &= (w(\{3, 1\}, \{\{2\}\}) - \text{MPW}_1(w_{-2})) = 6.5; \\ w_{-1}^{HM, \varphi}(\{23\}, \{\emptyset\}) &= w(\{1, 2, 3\}, \emptyset) - \text{MPW}_1(w) = 16.5. \end{aligned}$$

Indeed, solving this two-player game gives  $\text{MPW}_2(w_{-1}^{HM, \varphi}) = 8$  and  $\text{MPW}_3(w_{-1}^{HM, \varphi}) = 8.5$ , which coincides with the payoffs in the original game  $\text{MPW}_2(w)$  and  $\text{MPW}_3(w)$ , respectively.

It turns out that the MPW solution can be characterized in a manner similar to how the Shapley value is characterized using HM consistency.

**Theorem 9.** *The MPW solution, MPW, is the unique solution that satisfies 2-standardness ( $\mathbf{2S}^{\mathbf{X}}$ ) and HM consistency ( $\mathbf{HMC}^{\mathbf{X}}$ ).*

The proof relies on a different notion of HM consistency, which allows the removal of multiple players. More precisely, consider the following reduced TUX game  $w_{-T}^{HM,\varphi} \in \mathbb{W}(N \setminus T)$  defined by

$$w_{-T}^{HMC,\varphi}(S, \pi) = w(S \cup T, \pi) - \sum_{j \in T} \varphi_j(w_{-N \setminus (S \cup T)}) \quad (11)$$

for all  $(S, \pi) \in \mathcal{E}(N \setminus T)$ , where  $w_{-N \setminus (S \cup T)} \in \mathbb{W}(S \cup T)$  is defined in (5). The set version of HM consistency then reads as follows.

*Set HM Consistency,  $\mathbf{SHMC}^{\mathbf{X}}$ .* For all  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$ ,  $i \in N$ , and  $T \subseteq N \setminus \{i\}$ , we have  $\varphi_i(w) = \varphi_i(w_{-T}^{HMC,\varphi})$ .

Note that this boils down to HM consistency with  $T = \{j\}$ , i.e., it is obviously not a weaker requirement when allowing the removal of multiple players. Recall that for TU games, both notions of HM consistency—removing one or multiple players—are equivalent (Hart and Mas-Colell, 1989, Lemma 4.4). Leveraging the properties of the restriction given in (5), we can establish the same for TUX games.

**Proposition 10.** *A solution for TUX games  $\varphi$  satisfies HM Consistency ( $\mathbf{HMC}^{\mathbf{X}}$ ) if and only if  $\varphi$  satisfies Set HM Consistency ( $\mathbf{SHMC}^{\mathbf{X}}$ ).*

We finally mention that Dutta et al. (2010, Theorem 5) provide characterizations for alternative solutions for TUX games. These rely on notions of HM consistency that allow for the removal of multiple players at once (similar to  $\mathbf{SHMC}^{\mathbf{X}}$ ). It remains an open question whether an equivalence resembling Proposition 10 can be established for their properties and solutions.

## 5. Concluding Remarks

The widespread application of the Shapley value and its derivatives rest on its convincing characterizations, in particular, characterizations based on fairness properties such as balanced contributions by Myerson (1980) or the consistency properties due to Sobolev (1975) and Hart and Mas-Colell (1989). Often, solution concepts derived from the Shapley value are motivated by their derived analogon of the above properties in the specific setup. For generalizations of the Shapley value to games with externalities (TUX games), the literature is mainly focused on characterizations that derive from Shapley's original characterization based on additivity. While technically attractive, this property is less plausible from a normative perspective.

In this paper, we continued the work of Dutta et al. (2010) and Casajus et al. (2024), who investigate TUX games with changing player sets. We demonstrate that the balanced contributions property can be generalized to TUX games and yields a characterization of the MPW solution put forth by Macho-Stadler et al. (2007). Moreover, we introduce generalizations of Sobolev's consistency and of Hart and Mas-Colell's consistency to games with

externalities (TUX games). Again, this leads to consistency properties that are characteristic of the MPW solution.

Characterizations using consistency properties do not only allow us to distinguish competing solution concepts, but understanding the consistency property of the MPW solution further aides its computation and implementations via a mechanism. Moreover, it provides a template of consistency properties of allocation rules which derive from the MPW solution in specific applications. Similarly, the balanced contributions property in the presence of externalities is not only characteristic of the MPW solution, but it likely remains a plausible requirement for allocation rules that derive from the MPW solution in specific applications. In this regard, our work may pave the way for future research on allocation schemes in the presence of externalities for specific applications such as Saavedra-Nieves and Casas-Méndez (2023).

## 6. Acknowledgements

André Casajus: Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 388390901. Yukihiko Funaki: Funded by JSPS KAKENHI (18KK004, 22H00829).

## Appendix A. Appendix

First, we introduce further notation and insights before providing the proofs. For a given TUX game  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$ , we can compute an **auxiliary TU game**  $v_w^* \in \mathbb{V}(N)$ , in which each coalition  $S$  generates the worth generated by the grand coalition if we remove all other players  $N \setminus S$  from the game (Dutta et al., 2010). That is, the auxiliary TU game is defined by

$$v_w^*(S) = w_{-N \setminus S}(S, \emptyset) \quad \text{for all } S \subseteq N, \quad (\text{A.1})$$

where the restriction  $w_{-N \setminus S}$  is given in (5).

For  $N \subseteq \mathbf{U}$  and  $(T, \tau) \in \mathcal{E}(N)$ , the **scaled Dirac game with externalities**,  $\delta_{T, \tau} \in \mathbb{W}(N)$ , is defined by

$$\delta_{T, \tau}(S, \pi) = \begin{cases} \frac{1}{p_{N \setminus T}^*(\tau)}, & \text{if } (S, \pi) = (T, \tau), \\ 0, & \text{else,} \end{cases} \quad (\text{A.2})$$

for all  $(S, \pi) \in \mathcal{E}(N)$ . Clearly, these games generalize the Dirac TU games. For  $N \subseteq \mathbf{U}$  and  $T \subseteq N$ , the **Dirac TU game**  $\delta_T^N$  is defined by

$$\delta_T^N(S) = \begin{cases} 1, & \text{if } S = T, \\ 0, & \text{else} \end{cases} \quad (\text{A.3})$$

for all  $S \subseteq N$ . Note that every TUX game  $w \in \mathbb{W}(N)$  has a unique representation in terms of scaled Dirac games,

$$w = \sum_{(T, \tau) \in \mathcal{E}(N)} w(T, \tau) p_{N \setminus T}^*(\tau) \delta_{T, \tau}, \quad (\text{A.4})$$

i.e., the set  $\{(N, \delta_{T,\tau}) \mid (T, \tau) \in \mathcal{E}(N), T \neq \emptyset\}$  is a basis of the vector space of all TUX games with player set  $N$ .

#### Appendix A.1. Restriction of scaled Dirac games with externalities

For later convenience, we establish the following lemma (note that Casajus et al. (2024, Equation B.3) establish a similar result for unscaled Dirac games).

**Lemma 11.** *The restriction operator  $r^*$  is path independent and for all  $N \subseteq \mathbf{U}$ ,  $S \subseteq N$ , and  $(T, \tau) \in \mathcal{E}(N)$ , we have*

$$(\delta_{T,\tau})_{-S} = \begin{cases} \delta_{T,\tau_{-S}}, & \text{if } S \cap T = \emptyset, \\ \mathbf{0}^{N \setminus S}, & \text{otherwise.} \end{cases} \quad (\text{A.5})$$

*Proof of Lemma 11.* Applying the restriction (5) gives

$$(\delta_{T,\tau})_{-i}(S, \pi) \stackrel{(5)}{=} \frac{1}{n-s} \delta_{T,\tau}(S, \pi_{+i \sim \emptyset}) + \frac{|B|}{n-s} \sum_{B \in \pi} \delta_{T,\tau}(S, \pi_{+i \sim B}).$$

By definition of  $\delta_{T,\tau}$ , (A.5), the right-hand side of the upper equation vanishes unless  $(S, \pi_{+i \sim \emptyset}) = (T, \tau)$  or  $(S, \pi_{+i \sim B}) = (T, \tau)$ , for when it becomes  $\delta_{T,\tau}(T, \tau) / (n-t)$  or  $|B| \delta_{T,\tau}(T, \tau) / (n-t)$ , respectively. We get

$$(\delta_{T,\tau})_{-i}(S, \pi) \stackrel{(\text{A.5})}{=} \begin{cases} \frac{1}{n-t} \frac{1}{p_{N \setminus T}^*(\tau)}, & \text{if } i \in N \setminus T, (S, \pi) = (T, \tau_{-i}), \text{ and } \tau(i) = \{i\}, \\ \frac{|\tau(i)|}{n-t} \frac{1}{p_{N \setminus T}^*(\tau)}, & \text{if } i \in N \setminus T, (S, \pi) = (T, \tau_{-i}), \text{ and } |\tau(i)| > 1, \\ 0, & \text{if } i \in T. \end{cases}$$

By definition of  $p^*$ , (4),

$$\frac{1}{p_{(N \setminus \{i\}) \setminus T}^*(\tau_{-i})} = \begin{cases} \frac{1}{n-t} \frac{1}{p_{N \setminus T}^*(\tau)}, & \text{if } i \in N \setminus T \text{ and } \tau(i) = \{i\}, \\ \frac{|\tau(i)|}{n-t} \frac{1}{p_{N \setminus T}^*(\tau)}, & \text{if } i \in N \setminus T \text{ and } |\tau(i)| > 1. \end{cases}$$

Therefore, if  $i \in N \setminus T$ , then  $(\delta_{T,\tau})_{-i}(S, \pi) = \delta_{T,\tau_{-i}}(S, \pi)$ ; and if  $i \in T$ , then  $(\delta_{T,\tau})_{-i}(S, \pi) = 0$ . This means we obtain

$$\delta_{T,\tau} = \begin{cases} \delta_{T,\tau_{-i}}, & \text{if } i \in N \setminus T, \\ \mathbf{0}^{N \setminus \{i\}}, & \text{if } i \in T. \end{cases} \quad (\text{A.6})$$

Notice that  $((\delta_{T,\tau})_{-i})_{-j} = \mathbf{0}^{N \setminus \{i,j\}}$  if  $i \in T$  or  $j \in T$ ; and that  $((\delta_{T,\tau})_{-i})_{-j} = \delta_{T,\tau_{-\{i,j\}}} = ((\delta_{T,\tau})_{-j})_{-i}$  if  $i, j \in N \setminus T$ . Hence, restriction operator  $r^*$  is path independent for scaled Dirac games. Since these constitute a basis of the space of TUX games and since the restriction operator  $r^*$  is linear, this holds true for all TUX games (see Casajus et al. (2024, Theorem 7) for an alternative proof using unscaled Dirac games). In particular, the order in which players are removed from  $\delta_{T,\tau}$  does not matter, the removal of a coalition  $S \subseteq N$  is well-defined, and the claim follows from repeated application of (A.6).

### Appendix A.2. Average and auxiliary TU game of scaled Dirac games with externalities

We next establish the following Lemma (note that Casajus et al. (2024, Equation B.5) establish a similar result for unscaled Dirac games).

**Lemma 12.** *For all  $N \subseteq \mathbf{U}$  and  $(T, \tau) \in \mathcal{E}(N)$ , we have*

$$\bar{v}_{\delta_{T,\tau}} = \delta_T^N = v_{\delta_{T,\tau}}^* \in \mathbb{V}(N). \quad (\text{A.7})$$

*Proof of Lemma 12.* For  $N \subseteq \mathbf{U}$ ,  $(T, \tau) \in \mathcal{E}(N)$ , and  $S \subseteq N$ , we immediately get from the definition of the average game  $\bar{v}$ , (2), and the definition of the scaled Dirac games with externalities  $\delta_{T,\tau}$ , (A.2), that the factors  $p_{N \setminus T}^*(\tau)$  cancel out, and we have

$$\bar{v}_{\delta_{T,\tau}} \stackrel{(2),(4),(A.2)}{=} \delta_T^N.$$

For  $N \subseteq \mathbf{U}$  and  $(T, \tau) \in \mathcal{E}(N)$ , using the definition of the auxiliary game  $v^*$ , (A.1), the definition of the scaled Dirac games with externalities  $\delta_{T,\tau}$ , (A.2), Lemma 11, and  $p_\emptyset^*(\emptyset) = 1$  gives

$$\begin{aligned} v_{\delta_{T,\tau}}^*(S) &\stackrel{(\text{A.1})}{=} (\delta_{T,\tau})_{-N \setminus S}(S, \emptyset) \\ &\stackrel{\text{Lemma (11),(A.2)}}{=} \begin{cases} 1, & \text{if } T = S \text{ (and } \tau = \emptyset\text{),} \\ \mathbf{0}^{N \setminus S}, & \text{otherwise.} \end{cases} \\ &= \delta_T^N(S). \end{aligned}$$

**Remark 13.** *The lemma implies  $v_w^* = \bar{v}_w$  for all  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$ . This follows from linearity of both the average operator and the restriction operator as well as the uniqueness of the coefficients in (A.4). Further, we have*

$$\text{MPW}(w) \stackrel{(3)}{=} \text{Sh}(\bar{v}_w) = \text{Sh}(v_w^*) \quad \text{for all } w \in \mathbb{W}(N), N \subseteq \mathbf{U}.$$

### Appendix A.3. Proof of Lemma 5

The lemma follows from linearity of the involved operators, the fact that scaled Dirac games with externalities constitute a basis of  $\mathbb{W}$ , and from the following: if  $i \notin T$  then

$$(\bar{v}_{\delta_{T,\tau}})_{-i} \stackrel{(\text{A.7})}{=} (\delta_T^N)_{-i} = \delta_T^{N \setminus \{i\}} \stackrel{(\text{A.7})}{=} \bar{v}_{\delta_{T,\tau-i}} \stackrel{(\text{A.5})}{=} \bar{v}_{(\delta_{T,\tau})_{-i}},$$

whereas everything being the null game if  $i \in T$ .

### Appendix A.4. Proof of Theorem 4

*Existence:* We need to show that the MPW solution satisfies  $\mathbf{BC}^\mathbf{X}$ . By the definition of the MPW solution (3), i.e.,  $\text{MPW}_i(w) = \text{Sh}_i(\bar{v}_w)$ , and commutation of the average operator and of removal operator as shown in Lemma 5, i.e.,  $\bar{v}_{w-j} = (\bar{v}_w)_{-j}$ , we get

$$\text{MPW}_i(w_{-j}) \stackrel{(3)}{=} \text{Sh}_i(\bar{v}_{w_{-j}}) \stackrel{\text{Lemma 5}}{=} \text{Sh}_i((\bar{v}_w)_{-j}). \quad (\text{A.8})$$

Now we can confirm that the MPW solution inherits  $\mathbf{BC}^{\mathbf{X}}$  from the Shapley value,

$$\begin{aligned} \text{MPW}_i(w) - \text{MPW}_i(w_{-j}) &\stackrel{(A.8),(3)}{=} \text{Sh}_i(\bar{v}_w) - \text{Sh}_i((\bar{v}_w)_{-j}) \\ &\stackrel{\mathbf{BC} \text{ of Sh}}{=} \text{Sh}_j(\bar{v}_w) - \text{Sh}_j((\bar{v}_w)_{-i}) \stackrel{(A.8),(3)}{=} \text{MPW}_j(w) - \text{MPW}_j(w_{-i}). \end{aligned}$$

*Uniqueness:* Let  $\varphi$  satisfy and  $\mathbf{EF}^{\mathbf{X}}$  and  $\mathbf{BC}^{\mathbf{X}}$ . We show  $\varphi = \text{MPW}$  by induction on  $n$ . *Induction basis:* For  $n = 1$ , the claim  $\varphi = \text{MPW}$  is immediate from  $\mathbf{EF}^{\mathbf{X}}$ . *Induction hypothesis (IH):*  $\varphi(w) = \text{MPW}(w)$  for all  $w \in \mathbb{W}(N)$  such that  $n \leq \ell$ . *Induction step:* Let  $w \in \mathbb{W}(N)$  such that  $n = \ell + 1$ . By  $\mathbf{BC}^{\mathbf{X}}$ , we have

$$\varphi_i(w) - \varphi_j(w) = \varphi_i(w_{-j}) - \varphi_j(w_{-i}) \quad \text{for all } i, j \in N.$$

Summing up the equation over all  $j \in N$  gives

$$n\varphi_i(w) - \sum_{j \in N} \varphi_j(w) = \sum_{j \in N} (\varphi_i(w_{-j}) - \varphi_j(w_{-i})) \stackrel{IH}{=} \sum_{j \in N} (\text{MPW}_i(w_{-j}) - \text{MPW}_j(w_{-i})).$$

Finally, applying  $\mathbf{EF}^{\mathbf{X}}$  gives

$$n\varphi_i(w) = \sum_{j \in N} (\text{MPW}_i(w_{-j}) - \text{MPW}_j(w_{-i})) + w(N, \emptyset),$$

that is,  $\varphi_i(w)$  is uniquely determined for all  $i \in N$ . This implies  $\varphi_i(w) = \text{MPW}_i(w)$ .

#### Appendix A.5. Proof of Theorem 8

*Existence:* We need to show that the MPW solution satisfies  $\mathbf{SC}^{\mathbf{X}}$ . To this end, we compute the average game of the reduced game  $w_{-j}^{So, \text{MPW}}$ ,

$$\bar{v}_{w_{-j}^{So, \text{MPW}}}(S) \stackrel{(9),(2)}{=} \sum_{\pi \in \Pi(N \setminus S)} p^*(\pi) \left( \frac{s}{n-1} (w(S \cup \{j\}, \pi) - \text{MPW}_j(w)) + \frac{n-s-1}{n-1} w_{-j}(S, \pi) \right).$$

For  $(S, \pi) \in \mathcal{E}(N \setminus \{j\})$ , we obtain

$$\begin{aligned} \sum_{\pi \in \Pi(N \setminus S)} p^*(\pi) w(S \cup \{j\}, \pi) &\stackrel{(2)}{=} \bar{v}_w(S \cup \{j\}), \\ \sum_{\pi \in \Pi(N \setminus S)} p^*(\pi) \text{MPW}_j(w) &= \text{Sh}_j(\bar{v}_w), \text{ and} \\ \sum_{\pi \in \Pi(N \setminus S)} p^*(\pi) w_{-j}(S, \pi) &\stackrel{(2)}{=} \bar{v}_{w_{-j}}(S) \stackrel{\text{Lemma 5}}{=} (\bar{v}_w)_{-j}(S) = \bar{v}_w(S), \end{aligned}$$

where we used the definition of the average game  $\bar{v}_w$  (2); the fact that  $p^*$  is a probability distribution and the definition  $\text{MPW}_j(w) = \text{Sh}_j(\bar{v}_w)$  (3); and finally the commutation

Lemma 5. We get

$$\bar{v}_{w_{-j}^{So, \text{MPW}}} (S) = \frac{s}{n-1} (\bar{v}_w (S \cup \{j\}) - \text{Sh}_j (\bar{v}_w)) + \frac{n-s-1}{n-1} \bar{v}_w (S) = (\bar{v}_w)_{-j}^{So, \text{Sh}} (S). \quad (\text{A.9})$$

Applying the Shapley value on both sides and using Sobolev consistency (**SC**) of the Shapley value gives

$$\text{Sh}_i \left( \bar{v}_{w_{-j}^{So, \text{MPW}}} \right) \stackrel{(\text{A.9})}{=} \text{Sh}_i \left( (\bar{v}_w)_{-j}^{So, \text{Sh}} \right) \stackrel{\text{SC of Sh}}{=} \text{Sh}_i (\bar{v}_w)$$

for all  $i \in N \setminus j$ . By definition of MPW (3), the LHS is just  $\text{MPW}_i(w_{-j}^{So, \text{MPW}})$  and the RHS equals  $\text{MPW}_i(w)$ , which establishes the claim.

*Uniqueness:* Let the  $\varphi$  satisfy **EF<sup>X</sup>**, **2ST<sup>X</sup>**, and **SC<sup>X</sup>**. We show  $\varphi = \text{MPW}$  by induction on  $n$ . *Induction basis:* For  $n \leq 2$ , the claim is immediate from **EF<sup>X</sup>** and **2ST<sup>X</sup>**. *Induction hypothesis (IH):*  $\varphi(w) = \text{MPW}(w)$  for all  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$  such that  $n \leq \ell$ ,  $\ell \geq 2$ . *Induction step:* Let  $w \in \mathbb{W}(N)$  be such that  $n = \ell + 1$ . For  $i \in N$ , fix  $k \in N \setminus \{i\}$ . We obtain

$$\varphi_i (w) \stackrel{\text{SC}^X}{=} \varphi_i \left( w_{-k}^{So, \varphi} \right) \stackrel{IH}{=} \text{MPW}_i \left( w_{-k}^{So, \varphi} \right) \stackrel{\text{SC}^X}{=} \text{MPW}_i (w).$$

#### Appendix A.6. HM-reduction and player removal can be swapped

We prepare the proof of Proposition 10 by showing that it does not matter whether we first compute the HM-reduced game  $(w_{-T}^{HM, \varphi})_{-S}$  or whether we first remove players,  $(w_{-S})_{-T}^{HM, \varphi}$ .

**Lemma 14.** *For all TUX games  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$ , and coalitions  $S, T \in N$ , such that  $S \cap T = \emptyset$ , we have*

$$(w_{-T}^{HM, \varphi})_{-S} = (w_{-S})_{-T}^{HM, \varphi}.$$

*Proof of Lemma 14.* We first show that

$$w_{-S} (R, \rho) = \sum_{\tau \in \Pi(N \setminus R): \tau_{-S} = \rho} \frac{p_{N \setminus T}^*(\tau)}{p_{(N \setminus S) \setminus R}^*(\rho)} w(R, \tau) \quad (\text{A.10})$$

for all  $w \in \mathbb{W}(N)$ ,  $N \subseteq \mathbf{U}$ ,  $S \subseteq N$  and  $(R, \rho) \in \mathcal{E}(N \setminus S)$ . With the linearity of the restriction given in (5), we can exploit the representation of games by scaled Dirac games (A.4), that is, we have

$$w_{-S} = \sum_{(T, \tau) \in \mathcal{E}(N)} w(T, \tau) p_{N \setminus T}^*(\tau) (\delta_{T, \tau})_{-S}$$

for all  $w \in \mathbb{V}$  and  $S \subseteq N$ . Applying the formula for restrictions of scaled Dirac games (A.5) then gives

$$w_{-S} = \sum_{(T, \tau) \in \mathcal{E}(N): S \cap T = \emptyset} w(T, \tau) p_{N \setminus T}^*(\tau) \delta_{T, \tau_{-S}}.$$

The evaluation of  $w_{-S}$  at  $(R, \rho) \in \mathcal{E}(N \setminus S)$  vanishes if  $(R, \rho) \neq (T, \tau_{-S})$ . For  $(R, \rho) = (T, \tau_{-S})$ , it becomes  $1/p_{(N \setminus S) \setminus R}^*(\rho)$ , so that

$$w_{-S}(R, \rho) = \sum_{\tau \in \Pi(N \setminus R): \tau_{-S} = \rho} p_{N \setminus T}^*(\tau) w(R, \tau) \frac{1}{p_{(N \setminus S) \setminus R}^*(\rho)},$$

which reduces to the formula in the lemma by the definition of  $p^*$  (4).

Let now  $R \subseteq N \setminus (S \cup T)$  and  $\rho \in \Pi((N \setminus (S \cup T)) \setminus R)$ . We obtain

$$\begin{aligned} & \left( w_{-T}^{HM, \varphi} \right)_{-S}(R, \rho) \\ & \stackrel{(A.10)}{=} \sum_{\tau \in \Pi((N \setminus T) \setminus R): \tau_{-S} = \rho} \frac{p_{(N \setminus T) \setminus R}^*(\tau)}{p_{((N \setminus T) \setminus S) \setminus R}^*(\rho)} w_{-T}^{HM, \varphi}(R, \tau) \\ & \stackrel{(11)}{=} \sum_{\tau \in \Pi((N \setminus T) \setminus R): \tau_{-S} = \rho} \frac{p_{(N \setminus T) \setminus R}^*(\tau)}{p_{((N \setminus T) \setminus S) \setminus R}^*(\rho)} \left( w(R \cup T, \tau) - \sum_{j \in T} \varphi_j(w_{-N \setminus (R \cup T)}) \right). \end{aligned}$$

On the other hand, we have

$$\begin{aligned} & (w_{-S})_{-T}^{HM, \varphi}(R, \rho) \\ & \stackrel{(11)}{=} w_{-S}(R \cup T, \rho) - \sum_{j \in T} \varphi_j(w_{-N \setminus (R \cup T)}) \\ & \stackrel{(A.10)}{=} \sum_{\tau \in \Pi(N \setminus (R \cup T)): \tau_{-S} = \rho} \frac{p_{N \setminus (R \cup T)}^*(\tau)}{p_{(N \setminus S) \setminus (R \cup T)}^*(\rho)} \left( w(R \cup T, \tau) - \sum_{j \in T} \varphi_j(w_{-N \setminus (R \cup T)}) \right). \end{aligned}$$

Since  $R \subseteq N \setminus (S \cup T)$ , and since  $S$  and  $T$  are disjoint, both expressions are the same.

#### Appendix A.7. Proof of Proposition 10

It is clear that **SHMC<sup>X</sup>** implies **HMC<sup>X</sup>**. Let the solution  $\varphi$  satisfy **HMC<sup>X</sup>**. We show that  $\varphi$  satisfies **SHMC<sup>X</sup>** by induction on the size of the removed coalition  $T$ .

*Induction basis:* For  $t = 1$ , the claim is immediate by **HMC<sup>X</sup>**.

*Induction hypothesis (IH):* **SHMC<sup>X</sup>** implies **HMC<sup>X</sup>** for  $t = \ell$ .

*Induction step:* Let  $w \in \mathbb{W}(N)$ , and let  $T \subseteq N$  such that  $t = \ell + 1$ . For  $i \in N \setminus T$  and  $j \in T$ , we get

$$\varphi_i(v) \stackrel{IH}{=} \varphi_i \left( w_{-T \setminus \{j\}}^{HM, \varphi} \right) \stackrel{\mathbf{HMC}^X}{=} \varphi_i \left( \left( w_{-T \setminus \{j\}}^{HM, \varphi} \right)_{-j}^{HM, \varphi} \right).$$

Now, it suffices to show that  $(w_{-T \setminus \{j\}}^{HM, \varphi})_{-j}^{HM, \varphi}(S, \pi) = w_{-T}^{HM, \varphi}(S, \pi)$  for  $S \subseteq N \setminus T$  and  $\pi \in \Pi((N \setminus T) \setminus S)$  in order to establish **SHMC<sup>X</sup>**. Indeed, using the notation  $\tilde{S} =$

$(N \setminus (T \setminus \{j\})) \setminus (S \cup \{j\}) = N \setminus (S \cup T)$ , we get the following two equations:

$$\begin{aligned} (w_{-T \setminus \{j\}}^{HM,\varphi})_{-j}^{HM,\varphi}(S, \pi) &\stackrel{(11)}{=} w_{-T \setminus \{j\}}^{HM,\varphi}(S \cup \{j\}, \pi) - \varphi_j((w_{-T \setminus \{j\}}^{HM,\varphi})_{-\tilde{S}}) \\ &\stackrel{(11)}{=} w(S \cup T, \pi) - \sum_{k \in T \setminus \{j\}} \varphi_k(w_{-\tilde{S}}) - \varphi_j((w_{-T \setminus \{j\}}^{HM,\varphi})_{-\tilde{S}}); \end{aligned}$$

and

$$\begin{aligned} w_{-T}^{HM,\varphi}(S, \pi) &\stackrel{(11)}{=} w(S \cup T, \pi) - \sum_{k \in T \setminus \{j\}} \varphi_k(w_{-\tilde{S}}) - \varphi_j(w_{-\tilde{S}}) \\ &\stackrel{IH}{=} w(S \cup T, \pi) - \sum_{k \in T \setminus \{j\}} \varphi_k(w_{-\tilde{S}}) - \varphi_j((w_{-\tilde{S}})_{-T \setminus \{j\}}^{HM,\varphi}). \end{aligned}$$

Thus, we need to show that the order of removal and reduction does not matter, that is, we need to show

$$(w_{-T \setminus \{j\}}^{HM,\varphi})_{-\tilde{S}} = (w_{-\tilde{S}})_{-T \setminus \{j\}}^{HM,\varphi}.$$

Indeed, this is true because of Lemma 14. Note that we used the specifics of the restriction given in (5) only in the last step.

#### Appendix A.8. Proof of Theorem 9

*Existence:* It is well-known that MPW inherits **2ST<sup>X</sup>** from Sh. To establish **HMC<sup>X</sup>** of MPW, consider the average game of the reduced game  $w_{-j}^{HM,MPW}$ ,

$$\bar{v}_{w_{-j}^{HM,MPW}}(S) \stackrel{(10),(2)}{=} \sum_{\pi \in \Pi(N \setminus S)} p^*(\pi) (w(S \cup \{j\}, \pi) - \text{MPW}_j(w_{-N \setminus (S \cup \{j\})})).$$

for  $(S, \pi) \in \mathcal{E}(N \setminus \{j\})$ . We can insert

$$\begin{aligned} \sum_{\pi \in \Pi(N \setminus S)} p^*(\pi) w(S \cup \{j\}, \pi) &= \bar{v}_w(S \cup \{j\}), \text{ and} \\ \sum_{\pi \in \Pi(N \setminus S)} p^*(\pi) \text{MPW}_j(w_{-N \setminus (S \cup \{j\})}) &= \text{Sh}_j(\bar{v}_{w_{-N \setminus (S \cup \{j\})}}), \end{aligned}$$

giving

$$\bar{v}_{w_{-j}^{HM,MPW}}(S) = \bar{v}_w(S \cup \{j\}) - \text{Sh}_j(\bar{v}_{w_{-N \setminus (S \cup \{j\})}}) = (\bar{v}_w)_{-j}^{HM,Sh}(S) \quad (\text{A.11})$$

for all  $S \subseteq N \setminus \{j\}$ . Applying the Shapley value on both sides and using HM consistency (**HMC**) of the Shapley value gives

$$\text{Sh}_i(\bar{v}_{w_{-j}^{HM,MPW}}) \stackrel{(\text{A.11})}{=} \text{Sh}_i((\bar{v}_w)_{-j}^{HM,Sh}) \stackrel{\text{HMC of Sh}}{=} \text{Sh}_i(N, \bar{v}_w).$$

for all  $i \in N \setminus j$ . By definition of MPW (3), the LHS is just  $\text{MPW}_i(w_{-j}^{HM,\text{MPW}})$  and the RHS equals  $\text{MPW}_i(w)$ , which establishes the claim.

*Uniqueness:* Let  $\varphi$  satisfy **HMC<sup>X</sup>** and **2ST<sup>X</sup>**. By Proposition 10,  $\varphi$  satisfies **SHMC<sup>X</sup>**.

We first establish that  $\varphi$  is efficient (**EF<sup>X</sup>**). To this end, we proceed by induction on the size of the player set  $n$ . *Induction basis:* For  $n = 2$ , the claim is immediate from **2ST<sup>X</sup>**, and for  $n = 1$  it further follows from **SHMC<sup>X</sup>**. *Induction hypothesis (IH):* **SHMC<sup>X</sup>** and **2ST<sup>X</sup>** together imply **EF<sup>X</sup>** for  $n < \ell$ . *Induction step:* Let  $w \in \mathbb{W}(N)$  with  $n = \ell$ . Then, by definition of the reduced game (11) for  $T = N \setminus \{i\}$ , we have

$$w_{-N \setminus \{i\}}^{HM,\varphi}(\{i\}, \pi) = w(N, \pi) - \sum_{k \in N \setminus \{i\}} \varphi_k(w).$$

On the other hand, we have

$$w_{-N \setminus \{i\}}^{HM,\varphi}(\{i\}, \pi) \stackrel{\text{IH}}{=} \varphi_i(w_{-N \setminus \{i\}}^{HMC,\varphi}) \stackrel{\text{SHMC}^X}{=} \varphi_i(w).$$

The two equations together confirm **EF<sup>X</sup>**.

To establish uniqueness, i.e.,  $\varphi = \text{MPW}$ , we again proceed by induction on  $n$ . *Induction basis:* For  $n = 2$ , the claim is immediate from **2ST<sup>X</sup>**, and for  $n = 1$  it further follows from **EF<sup>X</sup>**. *Induction hypothesis (IH):*  $\varphi = \text{MPW}$  for  $n < \ell$ , where  $\ell > 2$ . *Induction step:* Let  $w \in \mathbb{W}(N)$  with  $n = \ell > 2$ . Fix two players  $i, k \in N$  and consider the reduced games on these two players  $w_{-N \setminus \{i,k\}}^{HM,\varphi}$  and  $w_{-N \setminus \{i,k\}}^{HM,\text{MPW}}$ . Note that

$$\begin{aligned} w_{-N \setminus \{i,k\}}^{HM,\varphi}(\{i\}, \pi) &= w(N \setminus \{k\}, \pi) - \sum_{j \in N \setminus \{i,k\}} \varphi_j(w_{-k}) \\ &\stackrel{\text{IH}}{=} w(N \setminus \{k\}, \pi) - \sum_{j \in N \setminus \{i,k\}} \text{MPW}_j(w_{-k}) = w_{-N \setminus \{i,k\}}^{HM,\text{MPW}}(\{i\}, \pi). \end{aligned}$$

Analogously,  $w_{-N \setminus \{i,k\}}^{HM,\varphi}(\{k\}, \pi) = w_{-N \setminus \{i,k\}}^{HM,\text{MPW}}(\{k\}, \pi)$ . By **2ST<sup>X</sup>**, the differences of payoffs of  $i$  and  $k$  must equal, i.e.,

$$\begin{aligned} &\varphi_i(w_{-N \setminus \{i,k\}}^{HM,\varphi}) - \varphi_k(w_{-N \setminus \{i,k\}}^{HM,\varphi}) \\ &\stackrel{\text{2ST}^X \text{ of } \varphi}{=} w_{-N \setminus \{i,k\}}^{HM,\varphi}(\{i\}, \pi) - w_{-N \setminus \{i,k\}}^{HM,\varphi}(\{k\}, \pi) \\ &= w_{-N \setminus \{i,k\}}^{HM,\text{MPW}}(\{i\}, \pi) - w_{-N \setminus \{i,k\}}^{HM,\text{MPW}}(\{k\}, \pi) \\ &\stackrel{\text{2ST}^X \text{ of } \text{MPW}}{=} \text{MPW}_i(w_{-N \setminus \{i,k\}}^{HM,\varphi}) - \text{MPW}_k(w_{-N \setminus \{i,k\}}^{HM,\varphi}). \end{aligned}$$

Applying **SHMC<sup>X</sup>** further augments this to the original game,

$$\varphi_i(w) - \varphi_k(w) = \text{MPW}_i(w) - \text{MPW}_k(w)$$

for arbitrary  $i, k \in N$ . Summing up this equation over all  $k \in N$  and applying  $\mathbf{EF}^X$  finally gives  $\varphi_i(w) = \text{MPW}_i(w)$  for all  $i \in N$ .

## References

- Albizuri, M.J., Arin, J., Rubio, J., 2005. An axiom system for a value for games in partition function form. *International Game Theory Review* 7, 63–73.
- Aldous, D., 1985. Exchangeability and related topics, in: Hennequin, P.L. (Ed.), Ecole d’été de probabilités de Saint-Flour, XIII—1983. Springer, Berlin. volume 1117 of *Lecture Notes in Mathematics*. chapter 1, pp. 1–198.
- Álvarez-Mozos, M., Ehlers, L., 2024. Externalities and the (pre)nucleolus in cooperative games. *Mathematical Social Sciences* 128, 10–15. doi:10.1016/j.mathsocsci.2024.01.003.
- Aumann, R., Maschler, M., 1985. Game Theoretic Analysis of a Bankruptcy Problem from the Talmud. *Journal of Economic Theory* 36, 195–213.
- Basso, F., Basso, L.J., Rönnqvist, M., Weintraub, A., 2021. Coalition formation in collaborative production and transportation with competing firms. *European Journal of Operational Research* 289, 569–581. doi:10.1016/j.ejor.2020.07.039.
- Bendel, D., Haviv, M., 2018. Cooperation and sharing costs in a tandem queueing network. *European Journal of Operational Research* 271, 926–933. doi:10.1016/j.ejor.2018.04.049.
- Bergantiños, G., Groba, C., Sartal, A., 2023. Applying the shapley value to the tuna fishery. *European Journal of Operational Research* 309, 306–318. doi:10.1016/j.ejor.2022.12.040.
- Bergantiños, G., Vidal-Puga, J., 2010. Realizing fair outcomes in minimum cost spanning tree problems through non-cooperative mechanisms. *European Journal of Operational Research* 201, 811–820. doi:10.1016/j.ejor.2009.04.003.
- Bergantiños, G., Moreno-Ternero, J.D., 2025. Revenue sharing at music streaming platforms. *Management Science* doi:<https://arxiv.org/abs/2310.11861>, arXiv:2310.11861.
- Bolger, E.M., 1989. A set of axioms for a value for partition function games. *International Journal of Game Theory* 18, 37–44.
- van den Brink, R., Chun, Y., 2011. Balanced consistency and balanced cost reduction for sequencing problems. *Social Choice and Welfare* 38, 519–529. doi:10.1007/s00355-011-0533-6.

- van den Brink, R., González-Arangüena, E., Manuel, C., del Pozo, M., 2014. Order monotonic solutions for generalized characteristic functions. *European Journal of Operational Research* 238, 786–796. doi:10.1016/j.ejor.2014.04.016.
- Brügemann, B., Gautier, P., Menzio, G., 2018. Intra firm bargaining and Shapley values. *The Review of Economic Studies* 86, 564–592.
- Casajus, A., 2014. Potential, value, and random partitions. *Economics Letters* 126, 164–166.
- Casajus, A., Funaki, Y., Huettner, F., 2024. Random partitions, potential, value, and externalities. *Games and Economic Behavior* 147, 88–106. doi:doi.org/10.1016/j.geb.2024.06.004.
- Casajus, A., Huettner, F., 2018. Decomposition of solutions and the Shapley value. *Games and Economic Behavior* 13, 1–23.
- Chambers, C.P., 2004. Consistency in the probabilistic assignment model. *Journal of Mathematical Economics* 40, 953–962. doi:10.1016/j.jmateco.2003.10.004.
- de Clippel, G., Serrano, R., 2008. Marginal contributions and externalities in the value. *Econometrica* 76, 1413–1436.
- Crane, H., 2016. The ubiquitous ewens sampling formula. *Statistical Science* 31, 1–19. doi:10.1214/15-sts529.
- Curiel, I., Pederzoli, G., Tijs, S., 1989. Sequencing games. *European Journal of Operational Research* 40, 344–351. doi:10.1016/0377-2217(89)90427-x.
- Davila-Pena, L., García-Jurado, I., Casas-Méndez, B., 2022. Assessment of the influence of features on a classification problem: An application to covid-19 patients. *European Journal of Operational Research* 299, 631–641. doi:10.1016/j.ejor.2021.09.027.
- Davis, M., Maschler, M., 1965. The kernel of a cooperative game. *Naval Research Logistics Quarterly* 12, 223–259. doi:10.1002/nav.3800120303.
- Driessen, T.S.H., 1991. A survey of consistency properties in cooperative game theory. *SIAM Review* 33, 43–59. doi:10.1137/1033003.
- Dutta, B., Ehlers, L., Kar, A., 2010. Externalities, potential, value and consistency. *Journal of Economic Theory* 145, 2380–2411.
- Dutta, B., Kar, A., 2004. Cost monotonicity, consistency and minimum cost spanning tree games. *Games and Economic Behavior* 48, 223–248. doi:10.1016/j.geb.2003.09.008.
- Ewens, W., 1972. The sampling theory of selectively neutral alleles. *Theoretical Population Biology* 3, 87–112.

- Feldman, B.E., 1996. Bargaining, Coalition Formation, and Value. Ph.D. thesis. State University of New York at Stony Brook.
- Fujinaka, Y., 2004. On the marginality principle in partition function form games. mimeo .
- González Arangüena, E., Manuel, C.M., del Pozo, M., 2015. Values of games with weighted graphs. European Journal of Operational Research 243, 248–257. doi:10.1016/j.ejor.2014.11.033.
- Gopalakrishnan, S., Granot, D., Granot, F., 2021. Consistent allocation of emission responsibility in fossil fuel supply chains. Management Science 67, 7637–7668. doi:10.1287/mnsc.2020.3874.
- Grabisch, M., Funaki, Y., 2012. A coalition formation value for games in partition function form. European Journal of Operational Research 221, 175–185. doi:10.1016/j.ejor.2012.02.036.
- Gul, F., 1989. Bargaining foundations of shapely value. Econometrica : journal of the Econometric Society 57, 81.
- Gómez-Rúa, M., Vidal-Puga, J., 2010. The axiomatic approach to three values in games with coalition structure. European Journal of Operational Research 207, 795–806. doi:10.1016/j.ejor.2010.05.014.
- Hafalir, I.E., 2007. Efficiency in coalition games with externalities. Games and Economic Behavior 61, 242–258. doi:10.1016/j.geb.2007.01.009.
- Hart, S., Mas-Colell, A., 1989. Potential, value, and consistency. Econometrica 57, 589–614.
- Ju, Y., Chun, Y., van den Brink, R., 2014. Auctioning and selling positions: A non-cooperative approach to queueing conflicts. Journal of Economic Theory 153, 33–45. doi:10.1016/j.jet.2014.05.007.
- Kamijo, Y., Kongo, T., 2012. Whose deletion does not affect your payoff? The difference between the Shapley value, the egalitarian value, the solidarity value, and the Banzhaf value. European Journal of Operational Research 216, 638–646.
- Kuipers, J., Mosquera, M.A., Zarzuelo, J.M., 2013. Sharing costs in highways: A game theoretic approach. European Journal of Operational Research 228, 158–168. doi:10.1016/j.ejor.2013.01.018.
- Lipovetsky, S., Conklin, M., 2001. Analysis of regression in game theory approach. Applied Stochastic Models in Business and Industry 17, 319–330.
- Littlechild, S.C., Owen, G., 1973. A simple expression for the shapely value in a special case. Management Science 20, 370–372.

- Lorenzo-Freire, S., Alonso-Mejide, J.M., Casas-Méndez, B., Hendrickx, R., 2007. Balanced contributions for tu games with awards and applications. European Journal of Operational Research 182, 958–964. doi:10.1016/j.ejor.2006.07.040.
- Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.I., 2020. From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence 2, 56–67.
- Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. Advances in neural information processing systems 30.
- Macho-Stadler, I., Pérez-Castrillo, D., Wettstein, D., 2007. Sharing the surplus: An extension of the Shapley value for environments with externalities. Journal of Economic Theory 135, 339–356.
- Maniquet, F., 2003. A characterization of the Shapley value in queueing problems. Journal of Economic Theory 109, 90–103.
- Maskin, E., 2003. Bargaining, coalitions and externalities, Econometric Society, Princeton. Presidential Address, Institute for Advanced Study.
- McQuillin, B., 2009. The extended and generalized Shapley value: Simultaneous consideration of coalitional externalities and coalitional structure. Journal of Economic Theory 144, 696–721.
- McQuillin, B., Sugden, R., 2016. Backward induction foundations of the shapley value. Econometrica 84, 2265–2280. doi:10.3982/ecta13191.
- McQuillin, B., Sugden, R., 2018. Balanced externalities and the Shapley value. Games and Economic Behavior 108, 81–92.
- Meng, F.Y., Gong, Z.W., Pedrycz, W., Chu, J.F., 2023. Selfish-dilemma consensus analysis for group decision making in the perspective of cooperative game theory. European Journal of Operational Research 308, 290–305. doi:10.1016/j.ejor.2022.12.019.
- Moulin, H., 2003. Fair division and collective welfare. MIT press.
- Myerson, R.B., 1977. Values for games in partition function form. International Journal of Game Theory 6, 23–31.
- Myerson, R.B., 1980. Conference structures and fair allocation rules. International Journal of Game Theory 9, 169–182.
- O'Neill, B., 1982. A problem of rights arbitration from the talmud. Mathematical Social Sciences 2, 345–371. doi:10.1016/0165-4896(82)90029-4.
- Peleg, B., Sudhölter, P., 2007. Introduction to the Theory of Cooperative Games. Springer-Verlag GmbH.

- Pham Do, K.H., Norde, H., 2007. The Shapley value for partition function form games. *International Game Theory Review* 9, 353–360.
- Pitman, J., 2006. Combinatorial Stochastic Processes. Springer.
- Potters, J., Sudhölter, P., 1999. Airport problems and consistent allocation rules. *Mathematical Social Sciences* 38, 83–102. doi:10.1016/s0165-4896(99)00004-9.
- Saavedra-Nieves, A., Casas-Méndez, B., 2023. On the centrality analysis of covert networks using games with externalities. *European Journal of Operational Research* 309, 1365–1378. doi:10.1016/j.ejor.2023.02.023.
- Shapley, L.S., 1953. A value for  $n$ -person games, in: Kuhn, H., Tucker, A. (Eds.), Contributions to the Theory of Games. Princeton University Press, Princeton. volume II, pp. 307–317.
- Shorrocks, A.F., 2012. Decomposition procedures for distributional analysis: A unified framework based on the Shapley value. *The Journal of Economic Inequality* 11, 99–126.
- Skibski, O., Michalak, T., 2019. Fair division in the presence of externalities. *International Journal of Game Theory* 49, 147–172. doi:10.1007/s00182-019-00682-4.
- Skibski, O., Michalak, T.P., Woolbridge, M., 2018. The stochastic Shapley value for coalitional games with externalities. *Games and Economic Behavior* 108, 65–80.
- Sobolev, A.I., 1975. The characterization of optimality principles in cooperative games by functional equations. Academy of Sciences of the Lithuanian SSR, Vilnius. volume 6. pp. 95–151.
- Sudhölter, P., Zarzuelo, J.M., 2017. Characterizations of highway toll pricing methods. *European Journal of Operational Research* 260, 161–170. doi:10.1016/j.ejor.2016.11.051.
- Thomson, W., 2015. Axiomatic and game-theoretic analysis of bankruptcy and taxation problems: An update. *Mathematical Social Sciences* 74, 41–59.
- Thomson, W., Velez, R.A., 2022. Consistent queueing rules. *Mathematical Programming* doi:10.1007/s10107-022-01905-5.
- Thrall, R.M., Lucas, W.F., 1963.  $n$ -person games in partition function form. *Naval Research Logistic Quarterly* 10, 281–293.
- Winter, E., 1992. The consistency and potential for values of games with coalition structure. *Games and Economic Behavior* 4, 132–144. doi:10.1016/0899-8256(92)90009-h.
- Xu, G., Driessen, T.S., Sun, H., Su, J., 2013. Consistency for the additive efficient normalization of semivalues. *European Journal of Operational Research* 224, 566–571. doi:10.1016/j.ejor.2012.08.018.

Yi, S.S., 1997. Stable coalition structures with externalities. Games and Economic Behavior 20, 201–237. doi:10.1006/game.1997.0567.

Young, H.P., 1985. Monotonic solutions of cooperative games. International Journal of Game Theory 14, 65–72.