

This is a repository copy of *The Balanced Threat Agreement for Individual Externality Negotiation Problems*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/161433/>

Version: Published Version

---

**Article:**

Borm, Peter, Funaki, Yukihiro and Ju, Yuan [orcid.org/0000-0002-7541-9856](https://orcid.org/0000-0002-7541-9856) (2020) The Balanced Threat Agreement for Individual Externality Negotiation Problems. *Homo Oeconomicus*. pp. 1-19. ISSN 2366-6161

<https://doi.org/10.1007/s41412-020-00097-7>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# *The Balanced Threat Agreement for Individual Externality Negotiation Problems*

**Peter Borm, Yukihiro Funaki & Yuan Ju**

**Homo Oeconomicus**

Journal of Behavioral and Institutional  
Economics

ISSN 0943-0180

Volume 37

Combined 1-2

Homo Oecon (2020) 37:67-85

DOI 10.1007/s41412-020-00097-7

**Your article is published under the Creative Commons Attribution license which allows users to read, copy, distribute and make derivative works, as long as the author of the original work is cited. You may self-archive this article on your own website, an institutional repository or funder's repository and make it publicly available immediately.**



# The Balanced Threat Agreement for Individual Externality Negotiation Problems

Peter Borm<sup>1</sup> · Yukihiro Funaki<sup>2</sup> · Yuan Ju<sup>3</sup>

Received: 27 November 2019 / Accepted: 30 April 2020 / Published online: 28 May 2020  
© The Author(s) 2020

## Abstract

This paper introduces a model to analyze individual externalities and the associated negotiation problem, which has been largely neglected in the game theoretic literature. Following an axiomatic perspective, we propose a solution, as a payoff sharing scheme, called the balanced threat agreement, for such problems. It highlights an agent's potential influences on all agents by threatening to enter or quit. We further study the solution by investigating its consistency. We also offer a discussion on the related stability issue.

**Keywords** Individual externalities · Negotiation · Balanced threat · Characterization · Consistency

**JEL Classification** C71 · D62 · D63 · D74

---

We thank the audience for their useful comments at the Economics Seminar of Ben-Gurion University of the Negev. We are grateful to the Guest Editor and the two anonymous reviewers for their detailed and constructive comments that substantially contribute to the improvement of the paper. Of course, all possible remaining errors are ours.

---

✉ Yuan Ju  
yuan.ju@york.ac.uk

Peter Borm  
P.E.M.Borm@uvt.nl

Yukihiro Funaki  
funaki@waseda.jp

<sup>1</sup> Department of Econometrics and Operations Research, Tilburg University, Tilburg, The Netherlands

<sup>2</sup> School of Political Science and Economics, Waseda University, Tokyo, Japan

<sup>3</sup> Department of Economics and Related Studies, University of York, York, UK

## 1 Introduction

Externalities arise whenever an (economic) agent undertakes an action that has an effect on another agent. It is called a negative externality when the effect turns to be a cost, or a positive externality when it is a benefit. An associated fundamental question in real life is how to resolve the externality-incurred conflicts among agents through negotiation.

By the source of externalities, generally, there may exist two types: coalitional externalities and individual externalities. The former has been well studied in the literature based on the model of *partition function form games*, proposed by Thrall and Lucas (1963). In partition function form games, a coalition may have different values when the coalition structures to which the coalition is included are different. That is, externalities come from the players' behavior of forming or breaking up coalitions. Various solution concepts have been proposed and analyzed by, among others, Myerson (1977), Bolger (1989), Feldman (1994), Maskin (2003), Macho-Stadler et al. (2007, 2010, 2018), Pham Do and Norde (2007), Ju (2007) and Borm et al. (2015).

Although individual externalities have been an important subject in economics, particularly in the context of institutional economics and the study of property rights (cf. Coase 1960), it has been largely overlooked by the game theoretic literature. Different from coalitional externalities, in many interactive environments, agents may not take any specifically competitive or collaborative actions against others. But just due to the fact that they may co-exist in a situation, each agent may affect other agents in the given situation. The impact generated from one agent to another agent may be caused by a particular action, or simply by the fact of existence of the agent, as a pub owner might like a fashion store to be its neighbor but not prefer a fire station.

The classic common-pool resource problem in economics well illustrates this (cf. Ostrom et al. 1994; Funaki and Yamato 1999). Consider a number of fishermen living around a lake and all have the right of fishing in it. Given the intrinsic differences among those agents in capabilities, skills, tools and equipments, etc., their activities would generate different impact on each other and result in different utilities. Groundwater exploitation by the farms in a certain area is of the same nature. According to Ostrom (1990, 2003), common-pool resources are often well governed by common property protocols, rather than through private property or state administration, which is a mechanism based on self-management by a local community in coordinating the users of the resources. This empirical finding suggests that through negotiation among the participating agents, but not by privatization of the resources, externality-based conflicts could be resolved. Indeed, this general negotiation problem is what we are interested to explore in the paper from a game theoretic perspective, though we would focus on a specific setting where individual agents may choose to enter or quit a certain situation, by which there appear individual externalities.

Ju and Borm (2008) introduced the model of *primeval games* to investigate individual externalities. Such an externality can happen because one agent might

have different values or utilities when the statuses of the surrounding agents are different. The framework of partition function form games does not model individual externalities as it assumes that all the players in the player set  $N$  are always present even if they do not form a coalition. The model of primeval games is different from the classical cooperative games in two main aspects: Primeval games do not consider cooperation (and, hence, the notion of coalition does not really apply), and primeval games take into account all situations in which only a subgroup of players (of the player set  $N$ ) are present. In this way, all possible externalities among players are modeled. Following a normative approach, Ju and Borm (2008) proposes a sequential negotiation process with which three compensation rules for primeval games are introduced. These compensation rules reflect the most prominent, yet different fundamental principles in the context, and hence, can serve as specific benchmarks to solve the associated compensation issue.

A restriction in the analysis of individual externalities in Ju and Borm (2008) is that the situation where all players appear is the only one in question but all the other situations only serve as reference points for the associated bargaining problem. In the current paper, we relax such a restriction and allow for analyzing the issue of bargaining and possible transfers with respect to an arbitrary group of agents. To this end, we extend the original notion of primeval games in Ju and Borm (2008) by accommodating the option for agents to have freedom to join in or leave the situation, and construct a model called *individual externality negotiation problem*.

In this paper, we are mostly interested in given a group of agents which payoff scheme(s) can be agreed upon by all agents, provided that an agents can choose to enter or quit at his or her own will. We first introduce the notion of threatening power in this environment. An agent's threatening power is her potential net influences on all agents' payoffs by executing either the option of entering the group or the option of quitting it. Following a unilateral perspective, we argue that an agreement can be reached when all agents have equal threatening power. This idea axiomatically motivates a new solution concept, called the *balanced threat agreement*. For any group of agents in an individual externality negotiation problem, we show that the balanced threat agreement exists and is unique. We provide a second characterization by means of local consistency.

The paper also addresses the stability issue of the balanced threat agreement. Since for any  $n$ -agents individual externality negotiation problem, there exists a balanced threat agreements for any given group of agents, it is interesting to analyze under which group its balanced threat agreement would make all agents have no incentive to change the situation. As a first attempt, we discuss the necessary and sufficient condition for the balanced threat agreement of the group  $N$  to be stable.

It is important to highlight that throughout the paper the notion of cooperation is *not* considered. As seen in the formal definitions below, when we introduce the notion of a group of agents, it is a neutral concept and does not imply cooperation among any agents of it at all. Neither the model nor the analysis considers cooperation in a coalition or between coalitions.

The paper has the following structure. The next section presents the general model: individual externality negotiation problem. Section 3 introduces two

straightforward axioms, efficiency and balanced threat, which axiomatically yields the solution concept of balanced threat agreement for this class of problems. Based on the local consistency, a second characterization is provided in Sect. 4. The final section concludes the paper by offering a discussion on the stability issue.

## 2 Individual Externality Negotiation Problems

Let  $N = \{1, \dots, n\}$  be the finite set of agents (or individuals). Note that for conventional convenience, we may also call these agents players, although strictly speaking the notion of a player here in the current model is different from a player in a classical transferable utility game. Here, agents are not concerned with the option of forming coalitions or not, as it is not the question in the context. A subset  $S$  of  $N$ , in order to be distinguished from the usual concept of *coalition* in cooperative games, is called a *group of individuals* (in short, a group  $S$ ). Here, the term of group is a neutral concept, which has nothing to do with cooperation or anything else, but simply means a set of individual agents in  $N$ .

A pair  $(i, S)$  that consists of an agent  $i$  and a group  $S$  of  $N$  to which  $i$  belongs is called an *embedded agent* (or *insider*) in  $S$ . Similarly, a pair  $(j, S)$  that consists of a player  $j$  and a group  $S$  of  $N$  to which  $j$  does not belong is called an *unembedded agent* (or *outsider*) with respect to  $S$ . Let  $\mathcal{E}(N)$  denote the set of embedded agents, i.e.

$$\mathcal{E}(N) = \{(i, S) \in N \times 2^N \mid i \in S\}.$$

Similarly,  $\mathcal{U}(N)$  denotes the set of unembedded agents, i.e.

$$\mathcal{U}(N) = \{(j, S) \in N \times 2^N \mid j \notin S\}.$$

**Definition 2.1** A mapping  $u$ , defined by

$$u|_{\mathcal{E}(N)} : \mathcal{E}(N) \longrightarrow \mathbb{R} \text{ and } u|_{\mathcal{U}(N)} : \mathcal{U}(N) \longrightarrow \{0\},$$

that assigns a real number  $u(i, S)$  to each embedded agent  $(i, S) \in \mathcal{E}(N)$  and  $u(j, S) = 0$  for all unembedded agents  $(j, S) \in \mathcal{U}(N)$  is an *individual-group function*. A tuple  $(N, u, M)$  where  $M \in 2^N \setminus \{\emptyset\}$  is called an *individual externality negotiation problem* with respect to group  $M$ . The set of individual externality negotiation problems with agent set  $N$  is denoted by  $IE^N$ .

The value  $u(i, S)$  represents the payoff, or utility, of agent  $i$ , given that all agents in  $S$  are present (or active) while all agents in  $N \setminus S$  are absent (or inactive). Thus, with different set of active agents, agent  $i$  may receive different payoffs, which suggests the externalities among agents. In case  $i$  is an unembedded agent, it means that agent  $i$  is not present, and by definition, all such agents have zero payoff. This is plausible

because if an agent does not appear in a certain situation, this agent receives neither benefit nor cost from the situation.<sup>1</sup> Moreover, we would like to emphasize that the agents in  $S$  *individually* come into view, without any cooperation among them. For any group of agents  $S \subseteq N$  and an individual-group function  $u$ , let  $u^S$  denote the vector  $(u(i, S))_{i \in N}$ .

Here, specifically, we use  $M$  to denote the group of agents that are assumed to be active. That is, for an individual externality negotiation problem  $(N, u, M) \in IE^N$ , the question we ask is which payoff sharing scheme can be agreed upon by all agents of  $N$  so that  $M$  is to be accepted as the group of individuals to be present while agents of  $N \setminus M$  remain outside, when all agents of  $N$  are allowed to take actions of either staying in or quitting  $M$  (if she was an insider of  $M$ ) and either staying outside or joining  $M$  (if she was an outsider with respect to  $M$ ).

**Definition 2.2** A solution, or an agreement, for  $IE^N$  is a function  $f$ , which associates with each individual externality negotiation problem  $(N, u, M)$  in  $IE^N$  a vector  $f(N, u, M)$  of individual payoffs in  $\mathbb{R}^N$ , i.e.,

$$f(N, u, M) = (f_i(N, u, M))_{i \in N} \in \mathbb{R}^N.$$

The following three-agent example<sup>2</sup> helps illustrate the class of individual externality negotiation problems.

**Example 2.3** A three-agent individual externality negotiation problem  $(N, u, M)$  with  $N = \{a, b, c\}$ ,  $M = \{a, b\}$  and the individual-group function  $u$  is given as follows. The underlying story could be well motivated by any common-pool resource situation, as explained in the introduction. Dependent which of the three agents would become active, i.e., as users of the resource (e.g., fishing in a pool, groundwater or underwater exploitation, coal mining, etc.), the agents may receive payoffs as specified above by  $u^S$ . Here, the very problem, as an example, is which payoff scheme can be agreed on by all the three agents so that they can accept that both agents  $a$  and  $b$  are active whereas  $c$  is absent.

$S$	$\emptyset$	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	$\{a, b, c\}$
$u^S$	(0, 0, 0)	(1, 0, 0)	(0, 1, 0)	(0, 0, 2)	(-1, 5, 0)	(-1, 0, 2)	(0, -2, 4)	(-2, 2, 3)

<sup>1</sup> It is equally plausible to consider a more general scenario where unembedded agents may obtain non-zero payoffs, which may reflect their “outside options” or the utilities of taking a different action from that of embedded agents. We assume zero payoffs for all unembedded agents as it helps the exposition of such a new model and its analysis, while it also has an advantage in certain contexts to clearly pin down a local analysis of the individual externalities confined to the given situation itself, without involving any possible impact from outside.

<sup>2</sup> To avoid the likely inconvenience in distinguishing an agent from the payoff, the agents are named as  $a, b$ , and  $c$  instead of 1, 2, 3 in the example.

### 3 The Balanced Threat Agreement

Consider the individual externality negotiation problem  $(N, u, M)$  given in Example 2.3. Let us start the analysis by discussing why simply allocating to the agents of  $N$  according to  $u^M$  may not necessarily be a desirable solution. Given now  $M = \{a, b\}$ , will the payoff vector  $(-1, 5, 0)$  to the three agents, respectively, be accepted? Obviously,  $a$  is unhappy with  $-1$  because if she chooses to quit, the new active group of agents will be  $\{b\}$  and thus  $a$  will get 0 which is better than a negative payoff. Meanwhile, her quitting will cause  $b$  to suffer from losing 4 ( $= 5 - 1$ ). Therefore,  $a$  would be able to demand more by threatening to quit. Agent  $b$  has no much support to demand 5 directly as in  $u^M$  because, apparently, if she chooses to quit, the corresponding new active group of agents will be  $\{a\}$  and then she will lost 5 but cause  $a$ 's situation to improve from  $-1$  to 1. Similarly, one can also check the behavior of the unembedded agent  $c$  as an outsider of the group  $M$ . In this case,  $c$  could ask more because by showing up he can get a higher payoff of 3 and at the same time make both  $a$  and  $b$  worse off. Hence, there seems to be a reason for  $a$  and/or  $b$  to pay  $c$  in order to keep him out of  $M$ . Consequently, the sharing scheme of  $(-1, 5, 0)$  is inapplicable.

From the above example, one can see that, for an individual externality negotiation problem, what matters for agents to demand more from a group is her potential influences to all agents by threatening to deviate from that status-quo (i.e., the given group). Accordingly, the final payoffs should be determined by a bargaining among all agents using their potential influences.

Now we formally explore this idea that will lead to a solution for individual externality negotiation problems.

Given a group  $M \in 2^N \setminus \{\emptyset\}$  and a corresponding payoff vector  $u^M$ . Any agent  $k \in N$  can make a demand  $x_k^M$  with respect to  $u^M$ , as no restriction is imposed on the freedom for agent  $k$  to stay with or leave  $M$  (if  $k$  is an insider of  $M$ ) or remain being out of or choose to join  $M$  (if  $k$  is an outsider of  $M$ ). We then have a demand vector  $x^M = (x_k^M)_{k \in N}$ . The question is how much an agent  $k$  should demand?

Firstly, a demand vector  $x^M$  can be accommodated if  $\sum_{k \in N} x_k^M \leq \sum_{i \in N} u_i^M$ . Thus, from an aggregate perspective, it is natural to introduce an axiom of efficiency towards a solution.

#### Axiom (Efficiency)

For any individual externality negotiation problem  $(N, u, M)$ , the payoff vector  $x^M$  is efficient if  $\sum_{k \in N} x_k^M = \sum_{i \in N} u_i^M$ .

Next we investigate how much an individual agent may demand from  $u^M$ . From a bargaining point of view and in the context of the *individual* externalities, all that an agent can do is to threaten the others by taking unilateral actions of either quitting  $M$  if  $i$  was an insider or joining  $M$  if she was an outsider, which will make impact on the rest of the agents due to the resulting externalities. Accordingly, an agent's

(individual) threatening power with respect to a demand vector  $x^M$ , denoted by  $\theta_i^M$ ,<sup>3</sup> can be measured as follows.

**Definition 3.1** (*Threatening power*) Given an individual externality negotiation problem  $(N, u, M)$ , for any player  $i \in N$ , the threatening power  $\theta_i^M$  is the total net influence that agent  $i$  can exert on all other agents by taking a unilateral action that deviates from the status-quo of  $M$ , i.e.,

$$\theta_i^M = \begin{cases} \sum_{k \in N \setminus \{i\}} (x_k^M - u_k^{M \setminus \{i\}}) - x_i^M & \text{if } i \in M; \\ \sum_{k \in N \setminus \{i\}} (x_k^M - u_k^{M \cup \{i\}}) - (x_i^M - u_i^{M \cup \{i\}}) & \text{if } i \in N \setminus M. \end{cases}$$

To help better understand this definition, we further explain it as follows. Consider an embedded agent  $i \in M$ . Supposing that all others’ demands  $(x_k^M)_{k \in N \setminus \{i\}}$  are known,  $i$  can make a demand  $x_i^M$  by threatening that if the demand is not satisfied, she will quit  $M$ , which is the only possible action for her to take in this case. If she quits,  $M$  is violated so that  $M \setminus \{i\}$  becomes a new group while each  $k \in M \setminus \{i\}$  will get  $u_k^{M \setminus \{i\}}$ , and  $i$  will get nothing since she now becomes an unembedded agent of  $M \setminus \{i\}$ . Hence, by taking the action of quitting,  $i$  can incur a damage on all other agents: They will lose  $\sum_{k \in N \setminus \{i\}} (x_k^M - u_k^{M \setminus \{i\}})$ , where  $u_k^{M \setminus \{i\}} = 0$  for all  $k \in N \setminus M$ . The cost for  $i$  to execute the action of quitting is  $x_i^M - 0$ .

Likewise, if  $i$  is an outsiders of  $M$ , she could demand  $x_i^M$  from  $M$  by threatening that, if her demand is not satisfied, she will join  $M$  to form a new group  $M \cup \{i\}$ . If she joins,  $M$  is violated so that within the new group  $M \cup \{i\}$  each  $k \in M$  will get  $u_k^{M \cup \{i\}}$ , and  $i$  will get  $u_i^{M \cup \{i\}}$  since she now becomes an insider of  $M \cup \{i\}$ . Hence, by making the action of joining,  $i$  can generate externalities to all other agents: They will lose  $\sum_{k \in N \setminus \{i\}} (x_k^M - u_k^{M \cup \{i\}})$ , where  $u_k^{M \cup \{i\}} = 0$  for all  $k \in N \setminus \{M \cup \{i\}\}$ . The cost for  $i$  to execute the action of joining is  $x_i^M - u_i^{M \cup \{i\}}$ .

**Example 3.2** Based on the three-agent individual externality negotiation problem  $(N, u, M)$  of Example 2.3, we illustrate the threatening powers of agents  $a, b$  and  $c$ , respectively. Given the demand vector  $x^{(a,b)} = \{x_a^{(a,b)}, x_b^{(a,b)}, x_c^{(a,b)}\}$ ,

$$\begin{aligned} \theta_a^{(a,b)} &= (x_b^{(a,b)} - 1) + (x_c^{(a,b)} - 0) - x_a^{(a,b)}; \\ \theta_b^{(a,b)} &= (x_a^{(a,b)} - 1) + (x_c^{(a,b)} - 0) - x_b^{(a,b)}; \\ \theta_c^{(a,b)} &= (x_a^{(a,b)} - (-2)) + (x_b^{(a,b)} - 2) - (x_c^{(a,b)} - 3). \end{aligned}$$

Since we assume that all agents are free to quit or join a group, each agent would try to negotiate for more by making use of her threatening power. It is then

<sup>3</sup> Precisely, the notation should be  $\theta_i^{x^M}$  as the threatening power is related to the demand vector  $x^M$ . For notational simplicity, where no confusion may arise, we use  $\theta_i^M$  instead of  $\theta_i^{x^M}$ .

plausible to expect that an agreement can only be accepted by all agents if the underlying payoff vector could make them have the same threatening power.

**Axiom (Balanced Threat)**

For any individual externality negotiation problem  $(N, u, M)$ , the payoff vector  $x^M$  satisfies balanced threat axiom if  $\theta_i^M = \theta_j^M$  for all  $i, j \in N$ .

It is interesting to see that there exists a unique solution that satisfies both the efficiency and the balanced threat axioms.

**Theorem 3.3** *For any individual externality negotiation problem  $(N, u, M)$ , there exists a unique solution that satisfies both the efficiency and the balanced threat axioms, which we call the balanced threat agreement  $\beta^M$ , and it is given by*

$$\beta_t^M = \frac{1}{n} \sum_{k \in M} u_k^M - \frac{1}{2} p_t + \frac{1}{2n} \sum_{k \in N} p_k \quad \text{for } t \in N,$$

where

$$p_t = \sum_{k \in M \setminus \{t\}} u_k^{M \setminus \{t\}} \quad \text{if } t \in M$$

$$p_t = \sum_{k \in M} u_k^{M \cup \{t\}} - u_t^{M \cup \{t\}} \quad \text{if } t \in N \setminus M.$$

**Proof** Let  $(N, u, M)$  be an individual externality negotiation problem. One can readily check that the balanced threat agreement defined above satisfies both the efficiency and the balanced threat axioms. To show that there exists a unique solution satisfying the two axioms, consider a solution  $x^M$  that satisfies both the efficiency and the balanced threat axioms. Below we show that  $x^M$  will necessarily be identical to  $\beta^M$ .

By efficiency, we know that

$$\sum_{k \in N} x_k^M = \sum_{i \in N} u_i^M,$$

and by the balanced threat axiom, we have

$$\text{For } i \in M, \quad \sum_{k \in M \setminus \{i\}} (x_k^M - u_k^{M \setminus \{i\}}) + \sum_{l \in N \setminus M} x_l^M - x_i^M = K,$$

$$\text{For } j \in N \setminus M, \quad \sum_{k \in M} (x_k^M - u_k^{M \cup \{j\}}) + \sum_{l \in N \setminus M \setminus \{j\}} x_l^M - (x_j^M - u_j^{M \cup \{j\}}) = K,$$

where  $K \in \mathbb{R}$ .

By using the first equation  $\sum_{k \in N} x_k^M = \sum_{i \in M} u_i^M$ , the second equation becomes

$$\sum_{k \in M} u_k^M - \sum_{k \in M \setminus \{i\}} u_k^{M \setminus \{i\}} - 2x_i^M = K,$$

and the third equation becomes

$$\sum_{k \in M} u_k^M - \sum_{k \in M} u_k^{M \cup \{j\}} + u_j^{M \cup \{j\}} - 2x_j^M = K.$$

Summing up all the second equations and the third equations, we have

$$\begin{aligned} n \sum_{k \in M} u_k^M - \sum_{i \in M} \sum_{k \in M \setminus \{i\}} u_k^{M \setminus \{i\}} - 2 \sum_{i \in M} x_i^M - \sum_{j \in N \setminus M} \sum_{k \in M} u_k^{M \cup \{j\}} + \sum_{j \in N \setminus M} u_j^{M \cup \{j\}} \\ - 2 \sum_{j \in N \setminus M} x_j^M = nK, \end{aligned}$$

where  $n = |N|$ . This and the first equation imply

$$K = \frac{n-2}{n} \sum_{k \in M} u_k^M - \frac{1}{n} \left( \sum_{i \in M} \sum_{k \in M \setminus \{i\}} u_k^{M \setminus \{i\}} + \sum_{j \in N \setminus M} \sum_{k \in M} u_k^{M \cup \{j\}} - \sum_{j \in N \setminus M} u_j^{M \cup \{j\}} \right)$$

This equation and the second equation imply

$$\begin{aligned} x_i^M = \frac{1}{n} \sum_{k \in M} u_k^M - \frac{1}{2} \sum_{k \in M \setminus \{i\}} u_k^{M \setminus \{i\}} + \frac{1}{2n} \left( \sum_{i \in M} \sum_{k \in M \setminus \{i\}} u_k^{M \setminus \{i\}} \right. \\ \left. + \sum_{j \in N \setminus M} \sum_{k \in M} u_k^{M \cup \{j\}} - \sum_{j \in N \setminus M} u_j^{M \cup \{j\}} \right) \end{aligned}$$

for  $i \in M$ . Similarly, the third equation implies

$$\begin{aligned} x_j^M = \frac{1}{n} \sum_{k \in M} u_k^M - \frac{1}{2} \left( \sum_{k \in M} u_k^{M \cup \{j\}} - u_j^{M \cup \{j\}} \right) + \frac{1}{2n} \left( \sum_{i \in M} \sum_{k \in M \setminus \{i\}} u_k^{M \setminus \{i\}} \right. \\ \left. + \sum_{j \in N \setminus M} \sum_{k \in M} u_k^{M \cup \{j\}} - \sum_{j \in N \setminus M} u_j^{M \cup \{j\}} \right) \end{aligned}$$

for  $j \in N \setminus M$ . This is equivalent to the following expression:

$$x_t^M = \frac{1}{n} \sum_{k \in M} u_k^M - \frac{1}{2} p_t + \frac{1}{2n} \sum_{k \in N} p_k.$$

Here

$$\begin{aligned} p_t &= \sum_{k \in M \setminus \{t\}} u_k^{M \setminus \{t\}} \quad \text{if } t \in M \\ p_t &= \sum_{k \in M} u_k^{M \cup \{t\}} - u_t^{M \cup \{t\}} \quad \text{if } t \in N \setminus M. \end{aligned}$$

Obviously,  $x^M = \beta^M$ . □

**Example 3.4** For the three-agent individual externality negotiation problem  $(N, u, M)$  of Example 2.3, together with the condition of efficiency that requires

$x_a^{\{a,b\}} + x_b^{\{a,b\}} + x_c^{\{a,b\}} = -1 + 5 + 0 = 4$ , one can readily compute that the balanced threat agreement  $\beta^{\{a,b\}} = \left\{ \frac{2}{3}, \frac{2}{3}, 2\frac{2}{3} \right\}$ .

It is useful to provide the balanced threat agreement  $\beta^N$  under  $N$  for an individual externality negotiation problem  $(N, u, N)$ . Without proof, we present the following result.

**Corollary 3.5** *For any individual externality negotiation problem  $(N, u, N)$ , the balanced threat agreement  $\beta^N$  is given by*

$$\beta_t^N = \frac{1}{n} \sum_{k \in N} u_k^N - \frac{1}{2} \left( \sum_{k \in N \setminus \{t\}} u_k^{N \setminus \{t\}} - \frac{1}{n} \sum_{k \in N} \sum_{l \in N \setminus \{k\}} u_l^{N \setminus \{k\}} \right),$$

for all  $t \in N$ .

This formula can be understood intuitively. Firstly, an agent  $t$  receives the equal division of the total payoff  $\sum_{k \in N} u_k^N$  under  $N$ . Then, it will be adjusted by the average of the difference between two forces.  $\sum_{k \in N \setminus \{t\}} u_k^{N \setminus \{t\}}$  is the aggregate payoff that all other agents can get if  $t$  deviates (i.e., leaves) from  $N$ .  $\frac{1}{n} \sum_{k \in N} \sum_{l \in N \setminus \{k\}} u_l^{N \setminus \{k\}}$  is the average aggregate payoff of all other agents if  $k$  deviates from  $N$ , or alternatively, is the aggregate payoff of all other agents if any agent may deviate from  $N$  with equal probability. We also remark that  $x^N$  only depends on the sum of utility vectors  $\sum_{k \in N} u_k^N$  and  $\sum_{l \in N \setminus \{k\}} u_l^{N \setminus \{k\}}$  for  $k \in N$ .

## 4 Consistency

In this section we study the underlying consistency property of the balanced threat agreement. Given an individual externality negotiation problem  $(N, u, M)$ . Now we consider a situation that an agent  $m$  in the set  $N$  leaves the problem forever. By leaving forever it would require that agent  $m$  will leave her benefit (or cost) behind and give up any possibility to exert any influence to other players of  $N$  but would expect to get some payoff in return. We then compare the solutions between the original situation (i.e., before  $m$  leaves) and the new situation (after  $m$  leaves).

The leaving agent  $m$  might have been a member of  $M$  or not. Let  $x^M$  be the solution of the original problem  $(N, u, M)$ . Suppose that when  $m$  leaves, she gets a payoff  $x_m^M$  according to the solution  $x^M$ , but she has to give up all other possible payoffs. This gives rise to the new situation as a *reduced* individual externality negotiation problem  $(N \setminus \{m\}, \hat{u}_x, \hat{M})$ .<sup>4</sup>

When  $m$  was in  $M$ , for the new problem  $\hat{u}$  we have the following condition

<sup>4</sup> For notational simplicity, when no confusion is caused, hereafter we will write  $\hat{u}$  instead of  $\hat{u}_x$  and write the new problem as  $(N \setminus \{m\}, \hat{u}, \hat{M})$ , though one should bear in mind that the  $\hat{u}$  is derived from  $(N, u, M)$  with respect to the specific solution  $x^M$  for  $(N, u, M)$ .

$$\sum_{i \in M \setminus \{m\}} \hat{u}_i^{M \setminus \{m\}} = \sum_{i \in M} u_i^M - x_m^M$$

as  $m$  has taken away  $x_m^M$  after she leaves. For any agent  $i \in M \setminus \{m\}$ ,  $\hat{u}$  is given by

$$\hat{u}_k^{M \setminus \{i, m\}} = u_k^{M \setminus \{i\}} + \frac{u_m^{M \setminus \{i\}}}{|M| - 2}, \quad \text{for all } k \in M \setminus \{i, m\}.$$

This is because after  $m$  leaves and takes away  $x_m^M$ , she also gives up any payoff in  $(N, u, M)$ . Thus, in case agent  $i$  deviates from  $M$ , everyone else of  $M$  should get the same benefit (or cost) from the leaving of  $m$ , and hence they evenly share the payoff  $u_m^{M \setminus \{i\}}$ . By the same token, for  $j \in N \setminus M$ ,  $\hat{u}$  is given by

$$\hat{u}_k^{M \cup \{j\} \setminus \{m\}} = u_k^{M \cup \{j\}} + \frac{u_m^{M \cup \{j\}}}{|M| - 1}, \quad \text{for all } k \in M \setminus \{m\},$$

and

$$\hat{u}_j^{M \cup \{j\} \setminus \{m\}} = u_j^{M \cup \{j\}}.$$

When  $m$  is not in  $M$ ,  $\hat{u}$  is given by

$$\sum_{i \in M} \hat{u}_i^M = \sum_{i \in M} u_i^M - x_m^M,$$

while for  $i \in M$ ,

$$\hat{u}_k^{M \setminus \{i\}} = u_k^{M \setminus \{i\}} \quad \text{for all } k \in M \setminus \{i\},$$

and for  $j \in (N \setminus M) \setminus \{m\}$ ,

$$\hat{u}_k^{M \cup \{j\} \setminus \{m\}} = u_k^{M \cup \{j\}} \quad \text{for all } k \in M \cup \{j\}.$$

It seems interesting to explore the possible consistency of a solution in terms of the reduced problem. Accordingly, we introduce the following axiom that requires the coincidence of the solutions between the two problems. While for the original problem, the solution is denoted by  $x^M$ , for the reduced problem, the notation of the solution would depend on the set of insiders. If the leaving agent  $m$  was in  $M$ , the solution is denoted by  $\hat{x}^{M \setminus \{m\}}$ , and if  $m$  is not in  $M$ , then it is denoted by  $\hat{x}^M$ .

**Axiom (Local Consistency)**

For any individual externality negotiation problem  $(N, u, M)$  and the reduced problem  $(N \setminus \{m\}, \hat{u}, \hat{M})$ , a solution  $x^M$  is locally consistent if for all  $k \in N \setminus \{m\}$ ,

$$x_k^M(N, u, M) = x_k^{\hat{M}}(N \setminus \{m\}, \hat{u}, \hat{M}).$$

**Theorem 4.1** *For any individual externality negotiation problem, the balanced threat agreement  $\beta^M$  under  $M$  satisfies local consistency.*

**Proof** First we will consider a case of  $m \in M$ . By the definition of  $\beta^M$ , for  $(N \setminus \{m\}, \hat{u}, \hat{M})$ , we have, for  $t \in N \setminus \{m\}$ ,

$$\begin{aligned} \beta_t^{M \setminus \{m\}}(N \setminus \{m\}, \hat{u}, \hat{M}) &= \frac{1}{n-1} \sum_{k \in M \setminus \{m\}} \hat{u}_k^{M \setminus \{m\}} - \frac{1}{2} \hat{p}_t + \frac{1}{2n} \sum_{k \in N \setminus \{m\}} \hat{p}_k \\ &= \frac{\sum_{k \in M} u_k^M - \beta_m^M}{n-1} - \frac{1}{2} \hat{p}_t + \frac{1}{2n} \sum_{k \in N \setminus \{m\}} \hat{p}_k, \end{aligned}$$

where

$$\begin{aligned} \hat{p}_t &= \sum_{k \in M \setminus \{m, t\}} \hat{u}_k^{M \setminus \{m, t\}} \quad \text{if } t \in M \setminus \{m\} \\ \hat{p}_t &= \sum_{k \in M \setminus \{m\}} \hat{u}_k^{M \setminus \{m\} \cup \{t\}} - \hat{u}_t^{M \setminus \{m\} \cup \{t\}} \quad \text{if } t \in N \setminus M. \end{aligned}$$

Then, for  $t \in M \setminus \{m\}$ , we have

$$\hat{p}_t = \sum_{k \in M \setminus \{m, t\}} \left( u_k^{M \setminus \{t\}} + \frac{u_m^{M \setminus \{t\}}}{|M| - 2} \right) = \sum_{k \in M \setminus \{m, t\}} u_k^{M \setminus \{t\}} + u_m^{M \setminus \{t\}} = \sum_{k \in M \setminus \{t\}} u_k^{M \setminus \{t\}} = p_t.$$

For  $t \in N \setminus M$ , we have

$$\begin{aligned} \hat{p}_t &= \sum_{k \in M \setminus \{m\}} \left( u_k^{M \cup \{t\}} + \frac{u_m^{M \cup \{t\}}}{|M| - 1} \right) - u_t^{M \cup \{t\}} \\ &= \sum_{k \in M \setminus \{m\}} u_k^{M \cup \{t\}} + u_m^{M \cup \{t\}} - u_t^{M \cup \{t\}} = \sum_{k \in M} u_k^{M \cup \{t\}} - u_t^{M \cup \{t\}} = p_t. \end{aligned}$$

These imply that, for  $t \in N \setminus \{m\}$ ,

$$\begin{aligned} \beta_t^{M \setminus \{m\}}(N \setminus \{m\}, \hat{u}, \hat{M}) &= \frac{\sum_{k \in M} u_k^M - \beta_m^M}{n-1} - \frac{1}{2} p_t + \frac{1}{2n} \sum_{k \in N \setminus \{m\}} p_k \\ &= \frac{\sum_{k \in M} u_k^M}{n-1} - \frac{1}{n-1} \left[ \frac{1}{n} \sum_{k \in M} u_k^M - \frac{1}{2} p_m + \frac{1}{2n} \sum_{k \in N} p_k \right] - \frac{1}{2} p_t + \frac{1}{2n} \sum_{k \in N \setminus \{m\}} p_k \\ &= \frac{1}{n-1} \left( 1 - \frac{1}{n} \right) \sum_{k \in M} u_k^M + \frac{1}{2(n-1)} p_m - \frac{1}{2n(n-1)} \sum_{k \in N} p_k - \frac{1}{2} p_t + \frac{1}{2(n-1)} \sum_{k \in N \setminus \{m\}} p_k \\ &= \frac{1}{n} \sum_{k \in M} u_k^M - \frac{1}{2n(n-1)} \sum_{k \in N} p_k - \frac{1}{2} p_t + \frac{1}{2(n-1)} \sum_{k \in N} p_k \\ &= \frac{1}{n} \sum_{k \in M} u_k^M + \frac{1}{2n(n-1)} \left( 1 - \frac{1}{n} \right) \sum_{k \in N} p_k - \frac{1}{2} p_t = \beta_t^M. \end{aligned}$$

This completes the proof for the case  $m \in M$ .

For the case of  $m \in N \setminus M$ , we remark that,

$$\hat{p}_t = \sum_{k \in M \setminus \{t\}} \hat{u}_k^{M \setminus \{t\}} = \sum_{k \in M \setminus \{t\}} u_k^{M \setminus \{t\}} \quad \text{for all } t \in M, \text{ and}$$

$$\hat{p}_t = \sum_{k \in M} \hat{u}_k^{M \cup \{t\}} - \hat{u}_t^{M \cup \{t\}} = \sum_{k \in M} u_k^{M \cup \{t\}} - u_t^{M \cup \{t\}} \quad \text{for all } t \in N \setminus M \setminus \{m\}.$$

These imply that for any  $t \in N \setminus \{m\}$

$$\begin{aligned} &\beta_t^M(N \setminus \{m\}, \hat{u}, \hat{M}) \\ &= \frac{1}{n-1} \sum_{k \in M} \hat{u}_k^M - \frac{1}{2} \hat{p}_t + \frac{1}{2(n-1)} \sum_{k \in N \setminus \{m\}} \hat{p}_k \\ &= \frac{\sum_{k \in M} u_k^M}{n-1} - \frac{\beta_m^M}{n-1} - \frac{1}{2} p_t + \frac{1}{2(n-1)} \sum_{k \in N \setminus \{m\}} p_k \\ &= \frac{\sum_{k \in M} u_k^M}{n-1} - \frac{1}{n-1} \left[ \frac{1}{n} \sum_{k \in M} u_k^M - \frac{1}{2} p_m + \frac{1}{2n} \sum_{k \in N} p_k \right] - \frac{1}{2} p_t + \frac{1}{2n} \sum_{k \in N \setminus \{m\}} p_k \\ &= \frac{1}{n-1} \left( 1 - \frac{1}{n} \right) \sum_{k \in M} u_k^M + \frac{1}{2(n-1)} p_m - \frac{1}{2n(n-1)} \sum_{k \in N} p_k \\ &\quad - \frac{1}{2} p_t + \frac{1}{2(n-1)} \sum_{k \in N \setminus \{m\}} p_k = \beta_t^M. \end{aligned}$$

This completes the proof for the case  $m \in N \setminus M$ . □

Next, consider an individual externality negotiation problem  $(N, u, M)$  where  $N = \{i, j\}$ . The following payoff vectors  $x^M$  are referred to as **standard** for two-agent individual externality negotiation problems.

If  $M = N = \{i, j\}$ , then

$$x_i^{\{i,j\}} = \frac{u_i^{\{i,j\}} + u_j^{\{i,j\}}}{2} + \frac{u_i^{\{i\}} - u_j^{\{j\}}}{4}, x_j^{\{i,j\}} = \frac{u_i^{\{i,j\}} + u_j^{\{i,j\}}}{2} + \frac{u_j^{\{j\}} - u_i^{\{i\}}}{4}.$$

If  $M = \{i\}, N \setminus S = \{j\}$ , then

$$x_i^{\{i\}} = \frac{u_i^{\{i\}}}{2} + \frac{u_i^{\{i,j\}} - u_j^{\{i,j\}}}{4}, x_j^{\{i\}} = \frac{u_i^{\{i\}}}{2} + \frac{u_j^{\{i,j\}} - u_i^{\{i,j\}}}{4}.$$

The standardness comes from a view of solving the problem by simply sharing the “surplus” equally, which is in the same spirit of the standard solution for transferable utility games. To see it clearly, taking  $x_i^{\{i,j\}}$  for example, it can be written as

$$x_i^{\{i,j\}} = \frac{u_i^{\{i\}} + 0}{2} + \frac{1}{2} \left( u_i^{\{i,j\}} + u_j^{\{i,j\}} - \frac{u_i^{\{i\}} + 0}{2} - \frac{u_j^{\{j\}} + 0}{2} \right).$$

That is, given agent  $i$ ’s individual payoff could be either  $u_i^{\{i\}}$  or 0, dependent on who the active agent is, it seems reasonable to take the average as the reference. So is for

agent  $j$ . Then, they share the surplus from  $u_i^{(ij)} + u_j^{(ij)}$  equally. Similarly, if  $M = \{i\}$ , we have

$$x_i^{(i)} = \frac{u_i^{(ij)} + 0}{2} + \frac{1}{2} \left( u_i^{(i)} - \frac{u_i^{(ij)} + 0}{2} - \frac{u_j^{(ij)} + 0}{2} \right).$$

Given agents now bargain over  $u_i^{(i)}$  as  $M = \{i\}$ , we take the average of payoffs in the other two scenarios as the reference and then share the surplus equally.

**Axiom (Standardness for 2-agent problems)**

For any 2-agent individual externality negotiation problem, a solution is standard if it yields the standard payoff vectors for any 2-agent individual externality negotiation problem.

**Theorem 4.2** *The balance threat agreement  $\beta^M$  is the only solution that satisfies local consistency and standardness for 2-agent problems.*

**Proof** Since the previous theorem shows that the balanced threat agreement satisfies local consistency, and it is easy to show that the solution satisfies the standardness for 2-agent problems, it suffices to show the uniqueness part.

For that purpose, we use an induction on the number of agents. It is apparent to see that for the case of 2-agents, any solution that is standard yields a unique payoff vector. Take any  $N$  such that  $|N| \geq 3$ . Next for any  $M \subseteq N$ , take any two solutions  $x^M, y^M$  which satisfy the local consistency. Then we have, if  $m \in M$ ,

$$x_t^M(N, u, M) = x_t^{M \setminus \{m\}}(N \setminus \{m\}, \hat{u}_x, M \setminus \{m\}), \quad \text{for all } t \in N \setminus \{m\},$$

and

$$y_t^M(N, u, M) = y_t^{M \setminus \{m\}}(N \setminus \{m\}, \hat{u}_y, M \setminus \{m\}), \quad \text{for all } t \in N \setminus \{m\};$$

and if  $m \in N \setminus M$ ,

$$x_t^M(N, u, M) = x_t^M(N \setminus \{m\}, \hat{u}_x, M), \quad \text{for all } t \in N \setminus \{m\}$$

and

$$y_t^M(N, u, M) = y_t^M(N \setminus \{m\}, \hat{u}_y, M) \quad \text{for all } t \in N \setminus \{m\}.$$

The induction hypothesis induces that, if  $m \in M$ ,

$$y_t^{M \setminus \{m\}}(N \setminus \{m\}, \hat{u}_y, M \setminus \{m\}) = x_t^{M \setminus \{m\}}(N \setminus \{m\}, \hat{u}_x, M \setminus \{m\}) \quad \text{for all } t \in N \setminus \{m\},$$

and if  $m \in N \setminus M$ ,

$$y_t^M(N \setminus \{m\}, \hat{u}_y, M) = x_t^M(N \setminus \{m\}, \hat{u}_x, M) \quad \text{for all } t \in N \setminus \{m\}.$$

Consider the case of  $m \in M$ . We have, for any  $t \in N \setminus \{m\}$ ,

$$\begin{aligned}
 &x_t^M(N, u, M) - y_t^M(N, u, M) \\
 &= x_t^{M \setminus \{m\}}(N \setminus \{m\}, \hat{u}_x, M \setminus \{m\}) - y_t^{M \setminus \{m\}}(N \setminus \{m\}, \hat{u}_y, M \setminus \{m\}) \\
 &= \left( \frac{1}{n-1} \sum_{k \in M \setminus \{m\}} \hat{u}_{x,k}^{M \setminus \{m\}} - \frac{1}{2} \hat{p}_t + \frac{1}{2(n-1)} \sum_{k \in N \setminus \{m\}} \hat{p}_k \right) \\
 &\quad - \left( \frac{1}{n-1} \sum_{k \in M \setminus \{m\}} \hat{u}_{y,k}^{M \setminus \{m\}} - \frac{1}{2} \hat{p}_t + \frac{1}{2(n-1)} \sum_{k \in N \setminus \{m\}} \hat{p}_k \right) \\
 &= \frac{1}{n-1} \sum_{k \in M} \hat{u}_{x,k}^M - \frac{1}{n-1} \sum_{k \in M} \hat{u}_{y,k}^M \\
 &= \frac{1}{n-1} \left( \sum_{k \in M} u_k^M - x_m^M \right) - \frac{1}{n-1} \left( \sum_{k \in M} u_k^M - y_m^M \right) \\
 &= \frac{1}{n-1} (y_m^M - x_m^M).
 \end{aligned}$$

By similar arguments, for the case of  $m \in N \setminus M$ , for any  $t \in N \setminus \{m\}$ , we have

$$\begin{aligned}
 &x_t^M(N, u, M) - y_t^M(N, u, M) \\
 &= x_t^M(N \setminus \{m\}, \hat{u}_x, M) - y_t^M(N \setminus \{m\}, \hat{u}_y, M) \\
 &= \left( \frac{1}{n-1} \sum_{k \in M} \hat{u}_{x,k}^M - \frac{1}{2} \hat{p}_t + \frac{1}{2(n-1)} \sum_{k \in N \setminus \{m\}} \hat{p}_k \right) \\
 &\quad - \left( \frac{1}{n-1} \sum_{k \in M} \hat{u}_{y,k}^M - \frac{1}{2} \hat{p}_t + \frac{1}{2(n-1)} \sum_{k \in N \setminus \{m\}} \hat{p}_k \right) \\
 &= \frac{1}{n-1} \sum_{k \in M} \hat{u}_{x,k}^M - \frac{1}{n-1} \sum_{k \in M} \hat{u}_{y,k}^M \\
 &= \frac{1}{n-1} \left( \sum_{k \in M} u_k^M - x_m^M \right) - \frac{1}{n-1} \left( \sum_{k \in M} u_k^M - y_m^M \right) \\
 &= \frac{1}{n-1} (y_m^M - x_m^M).
 \end{aligned}$$

These imply that for any  $m \in N$  and any  $t \in N \setminus \{m\}$ ,

$$x_t^M - y_t^M = \frac{-1}{n-1} (x_m^M - y_m^M).$$

Hence it also holds that for  $t \in N$  and  $m \in N \setminus \{t\}$ ,

$$x_m^M - y_m^M = \frac{-1}{n-1} (x_t^M - y_t^M).$$

These imply that for any  $m \in N$ ,

$$x_m^M - y_m^M = \left( \frac{-1}{n-1} \right)^2 (x_m^M - y_m^M).$$

Since  $|N| \geq 3$ , this implies that  $x_m^M = y_m^M$  for all  $m \in N$ . □

## 5 A Discussion on Stability

The main motivation of the paper is to propose a model to study a long neglected, yet general problem of individual externalities. Following a unilateral perspective, the balanced threat agreement, as a solution that provides a sensible payoff vector for any given group of agents, is axiomatically established. It is further characterized by means of consistency.

Since the solution we proposed focuses on a given group structure, one may argue that some agents might have incentive to deviate, if that would result in a new group structure and make them better off. This issue becomes more relevant if we consider the general  $n$ -agent environment, as for any  $M \in 2^N \setminus \{\emptyset\}$  there exists a balanced threat agreement. Thus, it makes sense to analyze under which group its balanced threat agreement would make all agents have no incentive to change the situation. As a first attempt and also for the fact that very often  $N$  is the focal point, we discuss the necessary and sufficient condition for the balanced threat agreement of the group  $N$  to be stable.

As we know, for any individual externality negotiation problem  $(N, u, M)$ , if  $i \in M$  leaves  $M$ , then she will get 0, and if  $j \in N \setminus M$  joins  $M$ , then he will get  $u_j^{M \cup \{j\}}$ . Accordingly, we can introduce the notion of stability as follows.

**Definition 5.1 (Stability)** For an individual externality negotiation problem  $(N, u, M)$ , a payoff vector  $x \in \mathbb{R}^N$  is called **stable** if it satisfies:

$$x_i \geq 0 \quad \text{for } i \in M, \quad x_j \geq u_j^{M \cup \{j\}} \quad \text{for } j \in N \setminus M.$$

The following theorem can be readily constructed by the definitions of efficiency and stability.

**Theorem 5.2** For an individual externality negotiation problem  $(N, u, M)$ , the stable payoff vector  $x \in \mathbb{R}^N$  which satisfies efficiency under  $M$  exists if and only if

$$\sum_{i \in M} u_i^M \geq \sum_{j \in N \setminus M} u_j^{M \cup \{j\}}$$

It is straightforward to see that, when  $M = N$ , any payoff vector  $x \in \mathbb{R}^N$  that satisfies  $x_k > 0$  for all  $k \in N$  is always stable.

The following theorem offers a necessary and sufficient condition for the balanced threat agreement  $\beta$  to be stable.

**Theorem 5.3** For an individual externality negotiation problem  $(N, u, N)$ , the balanced threat agreement  $\beta^N$  is stable if and only if

$$\frac{2 \sum_{k \in N} u_k^N}{n} \geq \max_{i \in N} \sum_{k \in N \setminus \{i\}} u_i^{N \setminus \{i\}} - \frac{\sum_{k \in N} \sum_{l \in N \setminus \{k\}} u_l^{N \setminus \{k\}}}{n}$$

**Proof** A necessary and sufficient condition of the stability of  $x^N$  is given by

$$\beta_i^N = \frac{1}{n} \sum_{k \in N} u_k^N - \frac{1}{2} \left( \sum_{k \in N \setminus \{i\}} u_k^{N \setminus \{i\}} - \frac{1}{n} \sum_{k \in N} \sum_{l \in N \setminus \{k\}} u_l^{N \setminus \{k\}} \right) \geq 0 \quad \text{for all } i \in N.$$

This is equivalent to

$$\frac{2 \sum_{k \in N} u_k^N}{n} \geq \sum_{k \in N \setminus \{i\}} u_k^{N \setminus \{i\}} - \frac{1}{n} \sum_{k \in N} \sum_{l \in N \setminus \{k\}} u_l^{N \setminus \{k\}} \quad \text{for all } i \in N.$$

Since  $\sum_{l \in N \setminus \{k\}} u_l^{N \setminus \{k\}}$  is independent on  $i$ , we have an equivalent expression:

$$\frac{2 \sum_{k \in N} u_k^N}{n} \geq \max_{i \in N} \sum_{k \in N \setminus \{i\}} u_k^{N \setminus \{i\}} - \frac{1}{n} \sum_{k \in N} \sum_{l \in N \setminus \{k\}} u_l^{N \setminus \{k\}}.$$

□

Apparently, a more general analysis of the stability issue seems challenging and requires further investigation, while alternative notions (e.g., collective versus unilateral, myopic versus farsighted) of stability are worth exploring.

### 6 Concluding Remarks

Given this is a new model about externalities that addresses a long-neglected problem, in addition to the aforementioned stability issue, naturally it could open up many other promising venues for future research: (1) We can explore alternative properties and characterizations of the balanced threat agreement. (2) One may adopt a strategic perspective to build non-cooperative bargaining protocols to analyze the individual externality negotiation problems. This will not only help discover the possible underlying strategic features of the balanced threat agreement in line with the research agenda of the Nash Program (cf. Trockel 2002), but also help motivate other plausible solution concepts. (3) We may expect to gain new insights on existing problems if applying the current model to concrete settings such as the river sharing problems (cf. van den Brink et al. 2018) where individual externalities prevail but are yet to be explicitly studied. (4) The current paper is mainly concerned with a specific negotiation problem with respect to a given active group of agents. It is natural and interesting to remove this specification but study the general negotiation issue with the entire set of agents, which no doubt will call for new solution concepts and extended analysis on stability problems.

(5) It may seem ambitious, and therefore quite challenging, but potentially very useful if we can introduce coalitional behavior into the modeling of the current problem. This would necessarily be a complicated model as it essentially combines the individual externality setting with partition function form games, but we may find it appealing in better fitting the real world.

## Compliance with Ethical Standards

**Conflict of interest** On behalf of all authors, the corresponding author states that there is no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bolger, E. M. (1989). A set of axioms for a value for partition function games. *International Journal of Game Theory*, 18, 37–44.
- Borm, P. E. M., Ju, Y., & Wettstein, D. (2015). Rational bargaining in games with coalitional externalities. *Journal of Economic Theory*, 157, 236–254.
- Coase, R. H. (1960). The problem of social cost. *Journal of Law and Economics*, 3, 1–44.
- Feldman, B. (1994). *Value with Externalities: the Value of the Coalitional Games in Partition Function Form*. Ph.D. Thesis. Stony Brook, NY: Department of Economics, State University of New York.
- Funaki, Y., & Yamato, T. (1999). The core of an economy with a common pool resource: a partition function form approach. *International Journal of Game Theory*, 28, 157–171.
- Ju, Y. (2007). The consensus value for games in partition function form. *International Game Theory Review*, 9, 437–452.
- Ju, Y., & Borm, P. E. M. (2008). Externalities and compensation: Primeval games and solutions. *Journal of Mathematical Economics*, 44, 367–382.
- 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.
- Macho-Stadler, I., Pérez-Castrillo, D., & Wettstein, D. (2010). Dividends and weighted values in games with externalities. *International Journal of Game Theory*, 39, 177–184.
- Macho-Stadler, I., Pérez-Castrillo, D., & Wettstein, D. (2018). Values for environments with externalities—The average approach. *Games and Economic Behavior*, 108, 49–64.
- Maskin, E. (2003). *Bargaining, Coalitions, and Externalities*. Econometric Society European Meeting, Stockholm, Sweden: Presidential Address to the Econometric Society.
- Myerson, R. (1977). Values for games in partition function form. *International Journal of Game Theory*, 6, 23–31.
- Ostrom, E. (1990). *Governing the Commons*. The Evolution of Institutions for Collective Action: Cambridge University Press. ISBN 0-521-40599-8.
- Ostrom, E. (2003). How types of goods and property rights jointly affect collective action. *Journal of Theoretical Politics*, 15, 239–270.
- Ostrom, E., Gardner, R., & Walker, J. (1994). *Rules, games, and common-pool resources*. University of Michigan Press, ISBN 978-0-472-06546-2.
- Pham Do, K. H., & Norde, H. (2007). The Shapley value for partition function form games. *International Game Theory Review*, 9, 353–360.

- Thrall, R. M., & Lucas, W. F. (1963).  $n$ -Person games in partition function form. *Naval Research Logistics Quarterly*, 10, 281–293.
- Trockel, W. (2002). Integrating the Nash program into mechanism theory. *Review of Economic Design*, 7, 27–43.
- van den Brink, R., He, S., & Huang, J. P. (2018). Polluted river problems and games with a permission structure. *Games and Economic Behavior*, 108, 182–205.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.