

## Power and welfare in bargaining for coalition structure formation

Shaheen Fatima<sup>1</sup> · Tomasz P. Michalak<sup>2,3</sup> ·  
Michael Wooldridge<sup>2</sup>

Published online: 14 September 2015  
© The Author(s) 2015

**Abstract** We investigate a noncooperative bargaining game for partitioning  $n$  agents into non-overlapping coalitions. The game has  $n$  time periods during which the players are called according to an *exogenous agenda* to propose offers. With probability  $\delta$ , the game ends during any time period  $t < n$ . If it does, the first  $t$  players on the agenda get a chance to propose but the others do not. Thus,  $\delta$  is a measure of the degree of democracy within the game (ranging from democracy for  $\delta = 0$ , through increasing levels of authoritarianism as  $\delta$  approaches 1, to dictatorship for  $\delta = 1$ ). We determine the subgame perfect equilibrium (SPE) and study how a player's position on the agenda affects his *bargaining power*. We analyze the relation between the *distribution of power* of individual players, the *level of democracy*, and the *welfare efficiency* of the game. We find that purely democratic games are welfare inefficient and that introducing a degree of authoritarianism into the game makes the distribution of power more equitable and also maximizes welfare. These results remain invariant under two types of player preferences: one where each player's preference is a total order on the space of possible coalition structures and the other where each player either likes or dislikes a coalition structure. Finally, we show that the SPE partition may or may not be *core* stable.

**Keywords** Bargaining · Coalition structure games · Externalities · Subgame perfect equilibrium · Core

---

Shaheen Fatima and Tomasz P. Michalak have contributed equally to this work.

---

✉ Shaheen Fatima  
s.s.fatima@lboro.ac.uk

Tomasz P. Michalak  
t.michalak@ox.ac.uk

Michael Wooldridge  
m.wooldridge@ox.ac.uk

<sup>1</sup> Department of Computer Science, University of Loughborough, Loughborough, Leicestershire, UK

<sup>2</sup> Department of Computer Science, University of Oxford, Oxford, UK

<sup>3</sup> Institute of Informatics, Warsaw University, Warszawa, Poland

## 1 Introduction

Many multi-agent systems require agents to effectively form coalitions in order to achieve goals that individuals cannot or to achieve them more efficiently [11]. The need to form coalitions arises not just in multi-agent settings but also in economic [19,35,45] and political [32] domains. Within this context, we focus on the following specific situation. There is a group of  $n$  agents who must decide how to partition themselves into non-overlapping coalitions. The players may have conflicting preferences over the possible coalition structures. Thus, the problem is to find a way to resolve such conflicts and form a partition that will be acceptable to all of them.

We take a non-cooperative bargaining approach to solve this problem. Table 1 provides a landscape of the existing literature on coalition structure generation. This can broadly be divided into two classes: *non-strategic* and *strategic*. Work within each of these two categories can be further divided in terms of the key problem addressed, the underlying assumptions, the notion of optimality, the notion of stability, the method of generating a coalition structure, and how the coalition structures are ranked (i.e., with or without externalities). If the agents have preferences not just over the coalitions to which they belong but over the entire coalition structure, then the preferences are said to have *externalities*.

In non-strategic models, the key problem is to generate a coalition structure that is optimal at a system-wide level. A key assumption that underlies these models is that a central entity has the power to impose such a socially optimal structure and that the individual players will accept it. As a consequence of this assumption, there is no notion of equilibrium or stability. The focus is only on the design of computationally efficient algorithms for generating a socially optimal structure. Although such an assumption simplifies the task of generating an optimal structure, it does not hold for multiagent systems. In a multiagent system, there is no central entity empowered to impose his decision on other agents. Each individual makes his own decision and has his own preferences that often conflict with others'. Consequently, it is important to study how conflicts can be resolved and stable structures formed through a process of strategic reasoning in which the individuals are the decision makers. Our research is aimed at achieving this goal.

Bargaining [35], a time-honoured mechanism for conflict resolution, has been widely studied in the game-theoretic literature on strategic decision making, and is particularly relevant to multiagent systems. In the strategic decision making models, the aim is to find coalition structures that are in *equilibrium* (i.e., are stable). The existing strategic models can be subdivided into *noncooperative* and *cooperative*. In the former, the process by which the individual agents bargain is explicitly incorporated in the model. In the latter, the bargaining protocol is abstracted away. See Sect. 6 for more details regarding the existing work in each category.

With reference to Table 1, our work belongs to the class of non-cooperative models with externalities. A number of non-cooperative bargaining protocols have been proposed in the literature [7–9,13] to resolve conflicts and facilitate coalition formation among self-interested agents. Many of them build upon Rubinstein's alternating offers game [40] and involve an *agenda* according to which the agents are called to make a new offer, accept or reject an existing offer, or make a counteroffer [7,8]. While many properties of such mechanisms have been extensively studied in the literature, there has been little work on the issue of how the position of an agent on the agenda influences his bargaining power<sup>1</sup> and how inequalities

<sup>1</sup> There are a number of alternative notions of power. For a detailed discussion of these and our own definition of power see Sect. 4.1.

**Table 1** A classification of existing models for optimal/ stable coalition structure generation

	Non-strategic decision making models	Strategic decision making models	
		Noncooperative models	Cooperative models
Key problem	To generate a socially optimal structure	To generate a stable structure	To generate a stable structure
Assumption	A central entity can impose a structure	No central entity individual agents decide upon a structure	A central entity can impose a structure
Notion of optimality	System level optimality	Individual level optimality	Individual level optimality
Notion of stability	None	Typically SPE or stationary equilibrium	Typically <i>core</i> if no externalities, and $\alpha$ , $\beta$ , $\gamma$ , or $\delta$ <i>core</i> otherwise
Method of generation	An algorithmic process	A bargaining protocol	Method not specified
Ranking of coalition structures	With/ without externalities	With/ without externalities	With/ without externalities

in the power of individual players affect system welfare. To date, the only such analysis of a bargaining mechanism for coalition structures was done by Bloch and Rottier [9] for simple preferences. (i.e., preferences where each player either likes or dislikes a coalition structure). However, in multi-agent settings, player preferences may be more general: partial or total orderings over the space of all possible coalition structures. It is therefore important to analyse bargaining in the context of such general player preferences.

In this article, we focus on coalition structure formation in a general setting that has a number of characteristics especially appropriate to multi-agent systems. In particular, we study the formation of multiple non-overlapping coalitions in situations with *externalities*. We assume that utility is not transferable, i.e, payoffs assigned to one agent cannot be assigned to another agent. Non-transferable utility appears, for instance, in settings where payoffs are expressed in terms of goals achievable by the agents [13]. Finally, we consider two types of preference orderings:

1. each player has unique rank for each possible coalition structure, and
2. each player associates one of two possible ranks with each possible coalition structure.

Our motivation for considering these two types of preferences is that we want to determine whether or not the results of our analysis are invariant to the way player preferences are modeled. A detailed discussion regarding the implications of this modeling choice is given in Sect. 4.6.

In order to model externalities and non-transferable utilities, we introduce *coalition structure games* (CSGs). CSGs encompass many important classes of coalitional games including hedonic games in which agents have preferences over coalitions they belong to, and NTU (non-transferable utility) games in partition function form that have been used to model multi-agent systems with various forms of externalities.

Using a CSG as the underlying game, we investigate a noncooperative bargaining game<sup>2</sup> that will allow a set of  $n$  players to decide how to partition themselves into coalitions.

<sup>2</sup> We will use the terms *protocol* and *game* interchangeably.

Noncooperative bargaining is conducted sequentially using an exogenously defined agenda. In this finite-horizon bargaining game with  $n$  discrete time periods, the players take it in turns to propose offers. An offer is a coalition structure, and the order in which the players are called to make offers is given by the agenda. An important element of our model is the parameter  $0 \leq \delta \leq 1$ , which represents the probability that bargaining ends in a given time period. Thus,  $\delta$  is a means to control the subset of agents that can make proposals during the process of bargaining and can be considered to be a measure of democracy within the game. For  $\delta = 1$ , only the first mover gets the chance to make a proposal—we call these *dictatorship games*, because a single agent is able to influence the outcome. For  $\delta = 0$ , all the  $n$  agents get a chance to make a proposal—we call these *democratic games*<sup>3</sup> in that every agent is able to influence the outcome. In between (i.e., for  $0 < \delta < 1$ ), we have games that are neither purely dictatorship nor purely democratic but have a degree of *authoritarianism*.

In a democratic game, although each agent is able to influence the outcome, the extents of their influence may differ. This extent is measured in terms of a player's *bargaining power*. A player's bargaining power crucially depends on his position on the agenda. Thus, we study the relationship between a player's position on the agenda and his bargaining power. We define a player's bargaining power as his ability to secure a favourable outcome. As an extreme example, for  $\delta = 1$ , the dictator always secures his most preferred coalition structure while all the other agents are completely powerless. However, for any  $\delta < 1$ , all agents on the agenda possess some degree of power, perhaps, however, not allocated evenly. We then analyse the relationship between the distribution of power, the level of democracy (embodied in  $\delta$ ), and the efficiency of the game, i.e., how far the bargained coalition structure is from the socially optimal coalition structure.

This article provides the first quantitative analysis of power and efficiency in the context of noncooperative bargaining for coalition structure formation. In the context of our bargaining model, the most important results, from both theoretical analysis and simulations, can be summarised as follows:

- We find that, assuming purely democratic games ( $\delta = 0$ ), the first mover has the lowest power and that power increases monotonically as we move from the first to the last position on the agenda. Surprisingly, however, the maximum is reached not for the last mover,  $n$ , but for the two last movers,  $n - 1$  and  $n$ .
- An even more surprising result is that an agent's average power (average taken over all possible player preference combinations) increases with the number of agents in the game. In fact, it is always profitable for incumbents to invite more agents to the game as long as the newcomers will be given the very first positions on the agenda.
- For  $\delta \approx 0$ , power monotonically increases from the first to the last player on the agenda. As  $\delta$  increases, the difference between the first and last mover's powers decreases and power becomes more balanced. With further increase in  $\delta$ , power monotonically decreases from the first to the last player on the agenda. In other words, as the degree of authoritarianism increases (i.e., as  $\delta$  approaches 1) there is a shift in power from the last movers toward the first movers.
- The above result is related to another notable finding related to the efficiency of democracy. Wittman [43,44] assumed that voters are rational, that elections are competitive, and that political transaction costs are low, and demonstrated that democracy is efficient (in the sense of Pareto optimality and wealth maximization). This is not so in our model. On the contrary, in our model, it is authoritarianism that maximizes the expected welfare of the system as a whole. We show that the reason behind this result is to be found in the

<sup>3</sup> See Sect. 2.3 for a detailed discussion on the interpretation of the term 'democracy'.

distribution of power. Democracy becomes inefficient if power is unevenly distributed, and some level of authoritarianism is needed to balance the power and maximize social welfare.

- The above results remain invariant under two types of player preferences: one where each player's preference is a total order on the space of possible coalition structures and the other where each player either likes or dislikes a coalition structure.

Bloch and Rottier [9] analyzed the relation between agenda and power and tried to answer questions such as: Can a head of state choose the composition of a cabinet by selecting the order in which party leaders attempt to form the government? Our setting is richer than that in Bloch and Rottier—one of its advantages is the ability to model externalities. Thus, our approach could be useful, for instance, to get insight into the distribution of power in structured multilateral negotiations on oligopolistic markets (such as armaments), where a government aims to distribute contracts among consortia of companies. Another important potential application is international negotiations on emission reduction. This is the setting with significant positive externalities—a reduction of emissions in one country also reduces pollution in other countries. Designing negotiations protocols for forming coalition structures has become a focal point of the entire body of literature [18,46]. It became especially important after the failure of Kyoto Protocol, where the grand coalition of all the countries failed to materialise. The literature recommended that subsequent protocols should be local and adjusted to the specific needs of the corresponding regions of the World [19]. In other words, the negotiation protocol should be focused on creating a coalition structure, not a single coalition.

The bargaining game studied in this article could be also used in multiagent applications that require automated formation of supply chains [17,30]. In this business setting, externalities play a significant role [24]. For example, Aerogistics<sup>4</sup> provided a platform for small-and medium-size manufacturers of aircraft components to form on-line supply-chain coalitions so that they were able to bid for manufacturing projects that were too large for them individually. The externalities in this setting arose from the fact that all aircraft components had to ultimately conform to the same standards. Consequently, the cost of standardization procedures incurred by any coalition depended on the number and structure of other winning coalitions.

The remainder of the article is organised as follows. We begin by introducing our model in Sect. 2. In Sect. 3, we provide subgame perfect equilibrium analysis for our bargaining game. In Sect. 4, we analyze, in various scenarios, the relation between the players' positions on the agenda and their power, and the welfare efficiency of the game. In Sect. 5, we show the relation between the noncooperative equilibrium and the  $\alpha$  core. This is followed by a discussion of related work in Sect. 6. Section 7 concludes. Appendix describes the Monte Carlo sampling procedure we used to conduct simulations.

## 2 The model

There are  $N = \{1, \dots, n\}$  players that want to decide how to partition themselves into non-overlapping coalitions. There are many possible ways in which this partitioning can be done and different players have different preferences over these possibilities. Each player wants the partition that he prefers the most, and the players are thus in a conflict about how to partition themselves.

<sup>4</sup> A British company located in Liverpool—see [www.aerogistics.com](http://www.aerogistics.com) for details.

In order to model the players' preferences over the possible coalition structures we define a *coalition structure game* and explore a noncooperative bargaining game for the players to resolve the conflict about how to partition themselves into a coalition structure. In Sect. 2.1, we will define 'coalition structure games' and in Sect. 2.2, the bargaining game.

## 2.1 Coalition structure game (CSG)

A *coalition structure*  $\pi$  is an exhaustive partition of  $N$  into disjoint coalitions; let  $\Pi(N)$  denote the set of all coalition structures over  $N$  and  $|\Pi(N)| = \text{Bell}(n)$  the number of possible coalition structures [37]. A coalition structure game  $\mathbb{G}$  is then defined by a tuple:

$$\mathbb{G} = \langle N, \succ_1, \dots, \succ_n \rangle,$$

where  $\succ_i \subseteq \Pi(N) \times \Pi(N)$  is a complete, non reflexive, and transitive preference relation for player  $i \in N$ , with the interpretation that if  $\pi_1 \succ_i \pi_2$ , then player  $i$  prefers coalition structure  $\pi_1$  more than the structure  $\pi_2$ . We will also consider a special type of game (briefly mentioned below, and dealt with in detail in Sect. 4.6) where each player either likes or dislikes a coalition structure.

Sometimes we find it useful to work with a numeric representation of preferences  $\succ_i$ , and for this reason, we now define the *rank*  $r_i(\pi)$  of a coalition structure  $\pi$  for an agent  $i$ . A player's most preferred coalition structure is the one with rank 1, then with rank 2, and so on. We will consider the following two kinds of rankings:

1. each player has a unique rank for each coalition structure, i.e., each  $r_i$  is a bijective map from  $\Pi(N)$  to  $\{1, \dots, \text{Bell}(n)\}$ , and
2. each  $r_i$  is a map from  $\Pi(N)$  to  $\{0, 1\}$  with  $r_i(\pi) = 1$  ( $r_i(\pi) = 0$ ) meaning that player  $i$  likes (dislikes)  $\pi$ .

Since the players may have conflicting preferences over the possible coalition structures, the bargaining game described in Sect. 2.2 is a mechanism for resolving this conflict.

## 2.2 The bargaining game

For a coalition structure game  $\mathbb{G}$ , we explore the following non-cooperative bargaining game for forming a coalition structure. This is a finite-horizon alternating offers game in which the players in  $N$  take it in turns to propose offers. The sequence in which the players are called to make offers is called the *bargaining agenda*.

**Definition 1** A bargaining agenda  $\rho$  for a game  $\mathbb{G}$  is a permutation of the first  $n$  integers where  $n = |N|$  is the number of players in  $\mathbb{G}$ .

In a nutshell, the rules of the game are as follows. Time is divided into discrete time periods. The game can run for at most  $n$  time periods. In any time period one of the players is chosen (as per the agenda) to propose an offer. All those who succeed the proposer on the agenda can then accept/ reject the offer. If all the successors accept, then the proposed structure is implemented and the game ends. Otherwise, we go to the next time period. There are two key elements of this bargaining game. One is with regard to the restriction imposed on those players whose offer is rejected: each such player must remain a singleton. The other element is the occurrence of a random event that can occur during any time period.

**Algorithm 1** Bargaining Game  $G$ **Require:** A  $\delta$  and a given agenda, i.e., an ordering of players in  $\rho$ 


---

```

1:  $OUT \leftarrow \emptyset$ ;
2:  $IN \leftarrow N$ ;
3:  $t \leftarrow 1$ ;
4: while  $t \leq n$  do
5:   Player  $\rho_t$  proposes a coalition structure  $\pi$  which is a partition of the players in  $IN$ .
     {A RandomEvent occurs with probability  $\delta$ . If it does, the game ends.}
6:   if (RandomEvent( $\delta$ )) then
7:     Go to Line 23
8:   end if
9:   for  $i = t + 1$  to  $n$  do
10:     $flag \leftarrow true$ 
11:    if (Player  $\rho_i$  rejects the proposal) then
12:       $OUT \leftarrow OUT \cup \{\rho_t\}$ 
13:       $IN \leftarrow IN - \{\rho_t\}$ 
14:       $flag \leftarrow false$ 
15:       $t \leftarrow t + 1$ 
16:      go to Line 4.
17:    end if
18:  end for
19:  if ( $flag == true$ ) then
20:    Go to Line 23
21:  end if
22: end while
23: The game ends and the coalition structure  $\pi \cup \{i : i \in OUT\}$  is implemented.

```

---

In more detail, the rules of the bargaining game are as follows. Bargaining starts in the first time period  $t = 1$  and proceeds as described in Algorithm 1. To begin, all the players in the game are in the set  $IN$ . The set  $OUT$  is initially empty (see Lines 1 and 2 in Algorithm 1). At  $t = 1$ , player Mover 1 makes an offer. An offer is a coalition structure  $\pi \in \Pi(IN)$ . After an offer is proposed, the game will end with probability  $\delta$ . With probability  $(1 - \delta)$  the game will continue to the next time period when Mover 2 will propose an offer (for details on the significance of  $\delta$  see Sect. 2.3).

In general, whether or not the game will end at  $t \geq 1$  is determined by a ‘RandomEvent’ (see Line 6 in Algorithm 1). If the random event does not occur, the game continues at Line 9. In the for loop of Line 9, the players  $\rho_{t+1}, \dots, \rho_n$  receive the offer. These players can accept or reject the current offer. If all these players accept, then the game ends and the structure  $\pi \cup \{i : i \in OUT\}$  is the outcome of the game. But if at least one of these players rejects, then the proposing player is moved from the set  $IN$  to the set  $OUT$ , time is incremented and we go to Line 4. Then, the next player on the agenda, i.e.,  $\rho_{t+1}$  proposes an offer. This process repeats until all the subsequent players (i.e.,  $\rho_{t+1}, \dots, \rho_n$ ) accept a proposal, or the time  $t = n$  is reached. If we reach  $t = n$ , the game ends and the structure  $\pi \cup \{i : i \in OUT\}$  is implemented.

Formally, the bargaining game is an  $n + 3$  tuple  $G = \langle N, \succ_1, \dots, \succ_n, \delta, \rho \rangle$ , however, for notational convenience, we will denote it  $G(n, \delta, \rho)$ ,  $G(n, \rho)$ ,  $G(n)$ ,  $G(\rho)$ , or just  $G$  when the other parameters are clear from context.

Let  $P$  denote the set of all possible preferences combinations for the  $n$  players. All other parameters remaining the same, we can obtain different bargaining games by varying  $\succ_1, \dots, \succ_n$ . Clearly, we have  $|P| = ((Bell(n))!)^n$  elements in the set  $P$  which means that we can have  $|P|$  possible bargaining games. Let  $\mathcal{G}$  denote the set of all these  $|P|$  games.



### 2.3 The significance of ‘RandomEvent’

We use ‘RandomEvent’ to model a wide spectrum of bargaining games. At one end of the spectrum, we have *dictatorship* games and at the other, we have *democratic* games. In between, we have games that are neither purely dictatorial nor purely democratic. For  $\delta = 1$ , only the first mover is given a chance to propose an offer. Once this offer is made, the game ends and no other player has any say in the outcome of the game. This is what we mean by a dictatorship game. For  $\delta = 0$ , all the  $n$  players are given a chance to propose an offer. Here,  $G$  is an  $n$  time period game. Since every player can exercise control over the outcome of the game, we call these democratic games. For  $0 < \delta < 1$ , the number of time periods in  $G$  will depend on  $\delta$  and will vary between 1 and  $n$ . Often, only the first few players on the agenda will get a chance to propose but the rest will not. These are *authoritarian* games.

Note that we use the term ‘democracy’ to refer to situations where all the players are eligible to propose. However, this does not necessarily mean that all of them will have equal bargaining power. Naturally, an ideal democracy would require that all players are eligible to vote and also that the power of their vote in the decision making process is equal. However, even in many political systems of the Western Hemisphere (such as Austria, Belgium, Finland, and The Netherlands), widely regarded as democratic, one could argue that not all votes are counted equally, at least, if we consider the entirety of the decision making process, i.e. from the act of voting to bills being passed in the parliament. This is because, while votes of individuals are counted equally, the parliaments are usually assembled using a not-entirely proportional rule such as the dHont method [20].

### 3 Subgame perfect equilibrium strategies

For a bargaining game  $G = \langle N, \succ_1, \dots, \succ_n, \delta, \rho \rangle$ , we will now show how to obtain equilibrium strategies. Since the bargaining game is a sequential moves game, the relevant equilibrium notion is subgame perfect equilibrium (SPE) [33]. A strategy profile is in subgame perfect equilibrium if it is a Nash equilibrium of every subgame of the bargaining game. We will first introduce notation and then characterize the SPE strategies. Let  $\pi_t^*(G)$  denote the equilibrium offer for time  $1 \leq t \leq n$ ; when  $G$  is evident from context, we will simply denote this offer as  $\pi_t^*$ . Also, let  $er_i(\pi_t)$  denote player  $i$ ’s expected rank from an offer  $\pi_t$  made at time  $t < n$ . Here,  $er_i(\pi_t)$  is defined as:

$$er_i(\pi_t) = \begin{cases} r_i(\pi_t) & \text{if } r_{\rho_k}(\pi_t) \leq r_{\rho_k}(\pi_{t+1}^*) \text{ for } k > t \\ \delta r_i(\pi_t) + (1 - \delta)r_i(\pi_{t+1}^*) & \text{otherwise} \end{cases}$$

For  $t = n$ , there is only one possible offer:  $\{\{i\} : i \in N\}$ . Thus, in the last time period, for each  $i \in N$ ,  $er_i(\{\{i\} : i \in N\}) = r_i(\{\{i\} : i \in N\})$ . In the following text,  $IN_t$  will denote the set of players in  $IN$  at time  $t$ . The set  $OUT_t$  will be analogous.

**Theorem 1** *For a given  $G$ , the following strategies form a subgame perfect Nash equilibrium. At  $t = n$ , the last mover will propose the structure  $\pi_n^* = \{\{i\} : i \in N\}$  and all the other players will accept. At  $t < n$ ,  $\rho_t$  will propose the structure  $\pi_t^* = \bar{\pi}_t \cup \{\{i\} : i \in OUT_t\}$  where*

$$\bar{\pi}_t \in \arg \min_{\pi \in \Pi(IN_t)} er_{\rho_t}(\pi) \text{ s.t. } er_{\rho_i}(\pi) \leq er_{\rho_i}(\pi_{t+1}^*) \text{ for } i > t.$$



**Table 2** The players' rankings for all possible coalition structures,  $N = \{1, 2, 3\}$ 

Game Coalition structure	(a)			(b)			(c)		
	Mover 1	Mover 2	Mover 3	Mover 1	Mover 2	Mover 3	Mover 1	Mover 2	Mover 3
$\{1, 2, 3\}$	1	2	1	1	3	2	1	2	1
$\{1, 2\}\{3\}$	4	4	4	2	2	1	4	4	4
$\{1, 3\}\{2\}$	3	3	3	3	1	3	3	3	3
$\{1\}\{2, 3\}$	5	1	2	4	4	4	5	1	5
$\{1\}\{2\}\{3\}$	2	5	5	5	5	5	2	5	2

Each player  $\rho_i$  ( $i > t$ ) will accept an offer  $\pi_t$  made at time  $t$  if  $er_{\rho_i}(\pi_t) \leq er_{\rho_i}(\pi_{t+1}^*)$ , and reject otherwise.

*Proof* The proof is based on backward induction. Consider the last time period:  $t = n$ . If bargaining proceeded to the last time period, there would be only one player (i.e.,  $\rho_n$ ) in  $IN_n$  and the first  $n - 1$  players would be in  $OUT_n$ . Thus, the only possible offer  $\rho_n$  could propose is  $\pi_n^* = \{\{i\} : i \in N\}$ . This would be accepted by all other players and the outcome of the game would be a structure comprised of  $n$  singleton coalitions.

Consider a time  $t < n$  and the following strategy: the offering player  $\rho_t$  will propose a structure that minimizes his expected rank subject to giving each subsequent player  $\rho_i$  ( $i > t$ ) an expected rank of at least  $er_{\rho_i}(\pi_{t+1}^*)$ . Since the players in  $OUT_t$  are singleton coalitions, this offer  $\pi_t^*$  will be:

$$\pi_t^* = \bar{\pi}_t \cup \{\{i\} : i \in OUT_t\}$$

where

$$\bar{\pi}_t \in \arg \min_{\pi \in \Pi(IN_t)} er_{\rho_t}(\pi) \text{ s.t. } er_{\rho_i}(\pi) \leq er_{\rho_i}(\pi_{t+1}^*) \text{ for } i > t.$$

Each player  $\rho_i$  ( $i > t$ ) will accept an offer  $\pi_t$  made at  $t$  if  $er_{\rho_i}(\pi_t) \leq er_{\rho_i}(\pi_{t+1}^*)$ , and reject otherwise.

Given these strategies for time  $t$ , those for the previous time period  $\pi_{t-1}^*$  will be:

$$\pi_{t-1}^* = \bar{\pi}_{t-1} \cup \{\{i\} : i \in OUT_{t-1}\}$$

where

$$\bar{\pi}_{t-1} \in \arg \min_{\pi \in \Pi(IN_{t-1})} er_{\rho_{t-1}}(\pi) \text{ s.t. } er_{\rho_i}(\pi) \leq er_{\rho_i}(\pi_t^*) \text{ for } i > t - 1.$$

Each player  $\rho_i$  ( $i > t - 1$ ) will accept an offer  $\pi_{t-1}$  made at  $t - 1$  if  $er_{\rho_i}(\pi_{t-1}) \leq er_{\rho_i}(\pi_t^*)$ , and reject otherwise.  $\square$

We will illustrate the application of the above defined equilibrium strategies by considering the following examples.

*Example 1* For a 3-player game, the preferences are as given in Table 2(a). Let  $\delta = 0$ . For the last time period, the equilibrium structure will be  $\{1\}\{2\}\{3\}$  which is least preferred by Mover 2 and Mover 3, and ranked 2 by Mover 1. For  $t = 2$ , the equilibrium structure will be  $\{1\}\{2, 3\}$  which is most preferred by Mover 2 and the second best in the ranking of Mover 3.

Thus, at  $t = 1$ , Mover 2 will not agree to any other coalition structure. The equilibrium for  $t = 1$  will be  $\{1\}\{2, 3\}$ . Here, Mover 2 secures his most preferred outcome.

*Example 2* Suppose the preferences are as given in Table 2(b). For  $t = n$ , the equilibrium structure will be  $\{1\}\{2\}\{3\}$  which is least preferred by each of the three players. For  $t = 2$ , the structure will be  $\{1\}\{2, 3\}$  which is ranked 4 by each player. Thus, at  $t = 1$ , Mover 1 will propose his most preferred structure  $\{1, 2, 3\}$ , and the others will accept because they both prefer it more than  $\{1\}\{2, 3\}$ . Here, Mover 1 is able to secure his most preferred outcome.

*Example 3* Suppose  $\delta = 0.2$  and the preferences are as given in Table 2(c). For  $t = n$ , the structure will be  $\{1\}\{2\}\{3\}$  which is least preferred by Mover 2, and ranked 2 by Mover 1 and Mover 3. For  $t = 2$ , Mover 2's equilibrium offer is the structure  $\{1\}\{2, 3\}$  since his expected rank from it is  $0.2 \times 1 + 0.8 \times 5 = 4.2$  while that from  $\{1\}\{2\}\{3\}$  is 5. For  $t = 1$ , Mover 1's equilibrium offer is his most preferred structure  $\{1, 2, 3\}$  as this gives him the least expected rank of  $0.2 \times 1 + 0.8 \times 5 = 4.2$ . Thus, if the random event occurs at  $t = 1$ , the resulting outcome will be  $\{1, 2, 3\}$ . Otherwise, Mover 2 will reject Mover 1's offer and propose  $\{1\}\{2, 3\}$ . If the random event occurs at  $t = 2$ , the resulting outcome will be  $\{1\}\{2, 3\}$ . Otherwise, Mover 3 will reject Mover 2's offer and the game will end at  $t = 3$  with  $\{1\}\{2, 3\}$  as the outcome.

*Example 4* For  $\delta = 0.2$ , consider the preferences in Table 2(b). For  $t = n$ , the structure will be  $\{1\}\{2\}\{3\}$  which is least preferred by all the players. For  $t = 2$ , Mover 2's equilibrium offer  $\{1\}\{2, 3\}$  will be accepted by Mover 1 and Mover 3. For  $t = 1$ , Mover 1's offer will be  $\{1, 2, 3\}$ . Whether or not the random event occurs, the game will end at  $t = 1$  with  $\{1, 2, 3\}$  as the outcome.

In Proposition 1, we characterize the time of agreement and, in Proposition 2, prove uniqueness of the SPE.

**Proposition 1** *If  $\delta = 0$  or  $\delta = 1$ , a game  $G$  will result in an immediate agreement, i.e., at  $t = 1$ . But for  $0 < \delta < 1$ , agreement can occur at any time  $1 \leq t \leq n$ .*

*Proof* If  $\delta = 0$ , with certainty, we have  $n$  time periods. Here, an offer made by the first mover in equilibrium for  $t = 1$  will be accepted by all the subsequent players.

If  $\delta = 1$ , with certainty, only the first mover makes an offer and the game ends at  $t = 1$ .

Next, consider the 3-player game with preferences given Table 2. We already demonstrated in Example 3 that an agreement can occur at  $t = 1$ ,  $t = 2$ , or  $t = 3$ .  $\square$

**Proposition 2** *The bargaining game has a unique SPE.*

*Proof* Since each player's preferences are given as a total order over the possible coalition structures, the equilibrium for each time period will be unique.  $\square$

In contrast, in Sect. 4.6, we study a special case where the player preferences are binary, that is, a coalition structure is either liked (1) or disliked (0) by a player. Then it will become clear that the uniqueness property will no longer hold.

## 4 Analysis of the relation between agenda position, power, and efficiency

In this section, we analyse how the power of a player is related to his position on the agenda and how the distribution of power in a mechanism affects its efficiency.

We start by defining the measures of power and efficiency in our setting (Sect. 4.1). Next, we describe the simulation setup and methodology (Sect. 4.2). We then move on to analyse power and efficiency for  $\delta = 0$ , i.e., under democracy (Sect. 4.3), and for  $\delta > 0$ , i.e., under authoritarianism (Sect. 2.3). Finally, we consider two more specific setups: (a) in Sect. 4.5, the bargaining game where a single player from  $N$  is *a priori* non-cooperative/cooperative; and (b) in Sect. 4.6, the bargaining game where players have binary preferences, that is, a coalition structure is either liked (rank 1) or disliked by a player (rank 0).

#### 4.1 Measures of power and efficiency

There are many different views on the nature of power and what power indices express or should express and what they measure. A popular view is that power index represents a *reasonable expectation* of a collective decision-making outcome [23]. Allingham [1] proposed that power be interpreted as a player's ability to influence outcomes irrespective of the *desirability* of these; the concept of power should therefore be independent of any concept of *utility*. In the same vein, Barry [5] distinguished between *success* and *power* by relying on specific player preferences. On the other hand, Roth [39] showed that, for risk-neutral players, the Shapley value is a von-Neumann-Morgenstern utility function. Roth [38], noted that the utility of playing a game from a given position is a reflection of the power which can be exerted by a player in that position. Likewise, Felsenthal and Machover [16] define a player's power as his expected payoff. We take the latter view and define a player's power as follows.

A player's power is a measure of his ability to secure his most preferred coalition structure as the subgame perfect equilibrium outcome of the bargaining game. For example, in Table 2(a), the first mover is only able to secure his least preferred structure while the second and third movers are able to secure their most preferred and second most preferred structures respectively. Thus, in this particular game, the second mover has the highest power, then the third mover, and then the first mover. Naturally, for other combinations of preferences, the first mover may have the highest power. This is the case in Table 2(b).

To account for all such possibilities, we measure the *power index* of the  $i$ 'th mover (i.e.,  $\rho_i$ ) in terms of his average power over all the games in  $\mathcal{G}$  (see Sect. 2.2 for a definition of  $\mathcal{G}$ ):

**Definition 2** The  $i$ 'th ( $i \in N$ ) mover's power index over the set of games in  $\mathcal{G}$  is defined as:

$$\mathbb{P}_{\rho_i}(\mathcal{G}) = 1 - \frac{\mathbb{E}_{\rho_i}(\mathcal{G}) - 1}{Bell(n) - 1}$$

where  $\mathbb{E}_{\rho_i}(\mathcal{G})$  denotes  $\rho_i$ 's average expected rank in the subgame perfect equilibrium for the games in  $\mathcal{G}$  and defined as follows:

$$\mathbb{E}_{\rho_i}(\mathcal{G}) = \frac{1}{(Bell(n)!)^n} \sum_{G \in \mathcal{G}} er_{\rho_i}(\pi_1^*(G)).$$

The following example illustrates the calculation of individual player power.

**Example 5** For  $N = \{1, 2, 3\}$ , we have  $|\Pi(N)| = 5$  and the number of all possible preference orderings equals  $(5!)^3 = 1,728,000$ . Let us consider the game, where  $\delta = 0$ . For 518,400 out of 1,728,000 preference orderings, the game ends with the coalition structure that is most preferred by Mover 1. In contrast, both Mover 2 and Mover 3 are able to secure their 1st choice for 704,040 preference orderings. The two latter movers are also able to secure also their 2nd choice more often, i.e. in 444,840 against 403,200 preference orderings. A detailed

**Table 3** (a) The numbers of preference orderings in which  $n = 3$  players can secure their  $i$ th ( $1 \leq i \leq 5$ ) most preferred structure

Number of preference orderings	(a)			(b)		
	Mover 1	Mover 2	Mover 3	Mover 1	Mover 2	Mover 3
Ranked 1	518,400	704,040	704,040	518,400	704,040	704,040
Ranked 2	403,200	444,840	444,840	806,400	889,680	889,680
Ranked 3	321,120	286,440	286,440	963,360	859,320	859,320
Ranked 4	263,520	181,800	181,800	1,054,080	727,200	727,200
Ranked 5	221,760	110,880	110,880	1,108,800	554,400	554,400
Total	1,728,000	1,728,000	1,728,000	4,451,040	3,734,640	3,734,640

(b) The numbers in (a) multiplied by the corresponding coalition structure rank

break-up of all 1, 728, 000 preference orderings is presented in Table 3(a). In Table 3(b), we report the numbers in part (a) multiplied by the corresponding rank of the coalition structure. For instance, for Mover 1, we have  $518,400 \times 1$ ,  $403,200 \times 2 = 806,400$ , and so on up to  $221,760 \times 5 = 1,108,800$ . By summing the entries in each column of Table 3(b) and dividing the result by the total number of preference orderings, we compute the players' average ranks:

$$\mathbb{E}_1(\mathcal{G}) = \frac{4,451,040}{1,728,000} = 2.57583, \quad \mathbb{E}_2(\mathcal{G}) = \mathbb{E}_3(\mathcal{G}) = \frac{3,734,640}{1,728,000} = 2.16125.$$

Next, we compute the players' powers:

$$\mathbb{P}_1(\mathcal{G}) = 1 - \frac{\mathbb{E}_1(\mathcal{G}) - 1}{\text{Bell}(3) - 1} = 0.606042,$$

and

$$\mathbb{P}_2(\mathcal{G}) = \mathbb{P}_3(\mathcal{G}) = 1 - \frac{\mathbb{E}_{2/3}(\mathcal{G}) - 1}{\text{Bell}(3) - 1} = 0.709688.$$

This concludes our example.

Clearly, the dictator in a game is a player  $i$  who can secure his most preferred structure irrespective of preferences of other players and will have  $\mathbb{P}_i(\mathcal{G}) = 1$ . On the other hand, a powerless player  $i$  has to accept any coalition structure proposed by others, irrespective of whether it suits its preferences or not and will have  $\mathbb{P}_i(\mathcal{G}) = 1/2$ . This is because his average rank, where average is taken over the equilibria for all possible combinations of player preferences, will have exactly the middle rank between 1 and  $\text{Bell}(n)$ . Thus, for  $\delta = 1$ , the first mover is the dictator while the others are powerless.

Power is important from an individual player's perspective. But it also important to understand how good the equilibrium outcome of a game is from the perspective of the society. For this, we use the notion of *social welfare*:

**Definition 3** The social welfare of a coalition structure is the sum of the individual players' ranks for the structure. The lower the sum, the higher the welfare.

**Definition 4** A welfare maximizing coalition structure is one that minimizes the sum of ranks of the players for the structure.

We will let  $\pi_{SW}(G) \in \Pi(N)$  denote a globally optimal coalition structure in game  $G$ .

A coalition structure bargained in  $G$  may not be the same as the socially optimal structure  $\pi_{SW}(G)$ . In order to measure how far a bargained coalition structure is from  $\pi_{SW}(G)$ , we define *efficiency ratio*.

**Definition 5** For a given  $n$  and  $\delta$ , the efficiency ratio,  $\mathcal{E}(G)$  for a game  $G$  is the ratio of the sum of the players' ranks for the globally optimal coalition structure for  $G$  and the sum of ranks for the bargained structure, i.e., we have:

$$\mathcal{E}(G) = \left( \sum_{i=1}^n r_i(\pi_{SW}(G)) \right) / \left( \sum_{i=1}^n er_i(\pi_1^*(G)) \right)$$

Since the sum of ranks for a bargained structure is never lower than that for the globally optimal structure, we have that  $\mathcal{E}(G) \leq 1$  for any  $n$  and  $\delta$ .

*Example 6* Let us consider the game in Table 2(a). In this particular preference ordering, the social welfare of  $\{1, 2, 3\}$ ,  $\{1, 2\}\{3\}$ ,  $\{1, 3\}\{2\}$ ,  $\{1\}\{2, 3\}$ , and  $\{1\}\{2\}\{3\}$  is  $1 + 2 + 1 = 4$ ,  $4 + 4 + 4 = 12$ ,  $3 + 3 + 3 = 9$ ,  $5 + 1 + 2 = 8$ , and  $2 + 5 + 5 = 12$ , respectively. Hence,  $\{1, 2, 3\}$  is the welfare maximizing coalition structure  $\pi_{SW}$ .

Having defined the measures of power and efficiency, let us describe in the next section our simulation setup and methodology.

## 4.2 Simulation setup and methodology

The model was implemented in C++ utilizing the message passing interface (MPI). This allowed for a very efficient parallel architecture. The computations were run on a cluster in which all computers were equipped with a 4-core AMD Opteron processor (2.0–2.3 Ghz) and 16 or 32 GB of RAM. At the height of the computations, 72 cores were utilized. Due to the nature of the model, we achieved a very efficient parallelization with an almost ideal linear speed-up.

For  $n = 3$ , it was possible to calculate the whole state space of preferences, that is,  $5!^3 = 1,728,000$  possible combinations for all 3 players (see Example 5). However, already for 4 players, the number of all possible combinations of preferences equals  $(15!)^4 = 2.9 \times 10^{48}$ . Consequently, for  $n \geq 4$ , we approximate  $\mathcal{P}_{\rho_i}(G)$  and  $\mathbb{E}_{\rho_i}(G)$  using Monte Carlo sampling. We generate a random combination of preference orderings from a uniform distribution. Since there is no bias in the original population, the samples should not be biased. In order to get unbiased samples, we use Knuth's shuffle algorithm [14]. Its details are presented in Appendix.

The quality of the approximation was assessed by comparing the average power of movers  $n - 1$  and  $n$ . When computed exactly, following Proposition 3, those powers should be the same. If they differed more than 1%, the number of Monte Carlo iterations was simply increased. In particular, for  $n = 4$ , at least 2,000,000 random preferences were sampled. For  $n = 9$ , we sampled 4,000,000 preferences per  $\delta$  and the simulation already took 24 h by utilizing 36 cores. While the largest game analyzed was for  $n = 14$  players and the program required approximately 64 GB of RAM and took over 20 min for only a single iteration, we typically report simulation results for up to 9 players as the precision obtained is within the assumed range.

*Example 7* As mentioned above, for  $n = 4$  we sampled at least 2,000,000 preference orderings—instances of the game. Then, we computed the following approximations of

players' powers:  $\mathbb{P}_1(\mathcal{G}) = 0.562431$ ,  $\mathbb{P}_2(\mathcal{G}) = 0.7061169$ , and  $\mathbb{P}_3(\mathcal{G}) = 0.7061180$ . Note that, following Proposition 3, the powers of movers 3 and 4 should be the same (if computed exactly). In this case, the powers of movers 3 and 4 differed only at the fourth decimal place (i.e. less than 0.0002 %). For  $n = 9$  and 4,000,000 sampled preference orderings, the power of the last two movers differed 1.07 %.

Having discussed our simulation setup we now turn to the discussion of power and efficiency.

### 4.3 Democracy ( $\delta = 0$ )

In this section we analyse power and efficiency for  $\delta = 0$ , i.e., under democracy.

**Power:** We can distinguish two types of power. In Example 1(a), player Mover 2 exercises his *power to reject* Mover 1's proposal. On the other hand, Mover 1 has the *power to propose* his best choice which does not have to necessarily correspond to the first choice of Mover 2 or subsequent movers. This indeed happens in Example 2(b), where Mover 1 proposes  $\{1, 2, 3\}$  and cannot be rejected because both other players like to cooperate with him.

From the definition of the game, it is easy to observe that, for  $n \geq 2$ , Mover 1 has only the power to propose, while Mover  $n$  has only the power to reject. For  $n = 2$ , Mover 1's power to propose will be equal to Mover 2's power to reject. To see why this is so, consider the number of preference orderings in which both movers can exercise their power. We have only two coalition structures  $\pi_1 = \{1, 2\}$  and  $\pi_2 = \{1\}\{2\}$  and four possible preference orderings:  $(1; 2)_1 = (\pi_1 \succ \pi_2; \pi_1 \succ \pi_2)$ ,  $(1; 2)_2 = (\pi_2 \succ \pi_1; \pi_2 \succ \pi_1)$ ,  $(1; 2)_3 = (\pi_1 \succ \pi_2; \pi_2 \succ \pi_1)$ , and  $(1; 2)_4 = (\pi_2 \succ \pi_1; \pi_1 \succ \pi_2)$ . Thus, we have:

$$\mathcal{G} = \{G(N, (1; 2)_1), G(N, (1; 2)_2), G(N, (1; 2)_3), G(N, (1; 2)_4)\}.$$

There is no conflict for the first two games and for each of these two games, the outcome is a structure that is most preferred (i.e., ranked 1) by both players. For  $G(N, (1; 2)_3)$ , Mover 2 has the power to reject Mover 1's offer in order to bring about his most preferred structure  $\{1\}\{2\}$ . Thus, the outcome of this game will be  $\{1\}\{2\}$  which is ranked 1 by Mover 2. Here, Mover 1 is powerless. The opposite happens for  $G(N, (1; 2)_4)$ . Now, Mover 1 will offer his most preferred structure  $\{1\}\{2\}$  and Mover 2 cannot gain anything by rejecting this offer. In other words, Mover 1 has the power to propose. Thus, the outcome of this game will be  $\{1, 2\}$  which is ranked 1 by Mover 1 and 2 by Mover 2. We therefore have  $\mathbb{E}_{\rho_1}(\mathcal{G}) = \mathbb{E}_{\rho_2}(\mathcal{G}) = 5/4$  and  $\mathbb{P}_{\rho_1}(\mathcal{G}) = \mathbb{P}_{\rho_2}(\mathcal{G}) = 3/4$ .

**Proposition 3** For  $n \geq 2$ , the last two movers have equal power, i.e.,  $\mathbb{P}_{\rho_{n-1}}(\mathcal{G}) = \mathbb{P}_{\rho_n}(\mathcal{G})$ .

*Proof* We recall that during the time period  $n - 1$  both Mover  $n - 1$  and Mover  $n$  have preference orderings over  $Bell(n)$  coalition structures but, as per the rules of bargaining, only two coalition structures are feasible:  $\pi = \{1\}\{2\} \dots \{n\}$  and  $\pi' = \{1\}\{2\} \dots \{n-2\}\{n-1, n\}$ . Since we consider all possible combinations of preference orderings, the number of times  $\pi \succ \pi'$  for Mover  $n - 1$  will be the same as the number of times  $\pi \succ \pi'$  for Mover  $n$ . Thus,  $\mathbb{P}_{\rho_{n-1}}(\mathcal{G}) = \mathbb{P}_{\rho_n}(\mathcal{G})$ .  $\square$

The fact that the last two movers have equal power is convenient when computing power in bigger games. Figure 1 shows the players' powers in games of size  $n = 2, 3, \dots, 9$ , from which we make the following observations:

- (a) For all  $n \geq 2$ , the power of Mover  $n$  is identical to the power of Mover  $n - 1$  (Proposition 3).



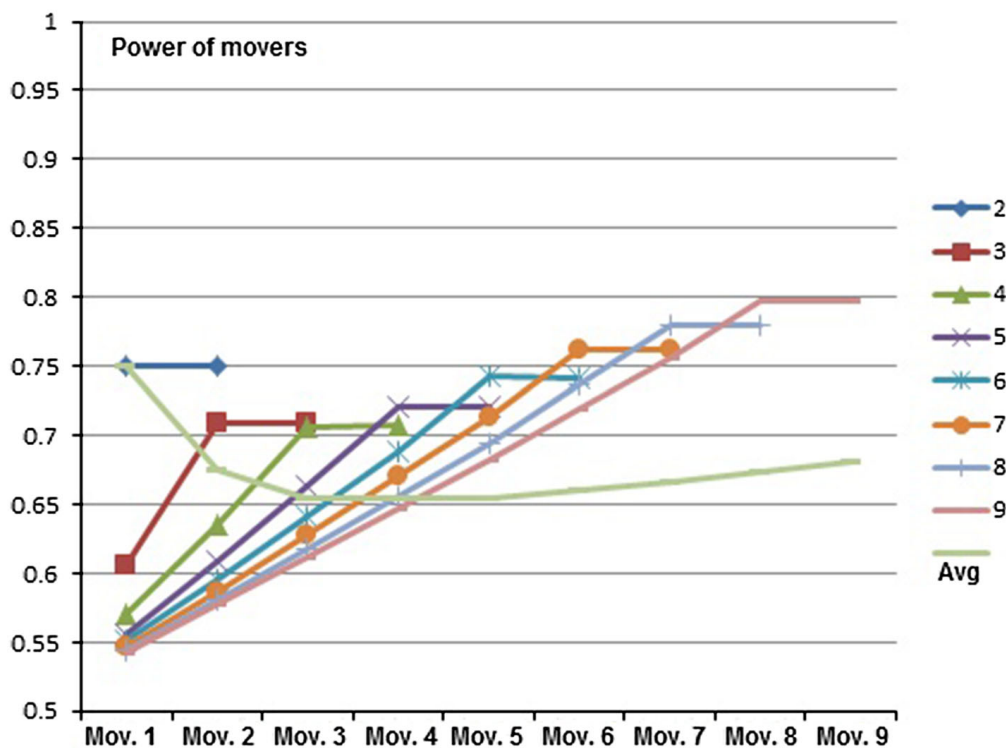


Fig. 1 The power in games of different size ( $\delta = 0$ )

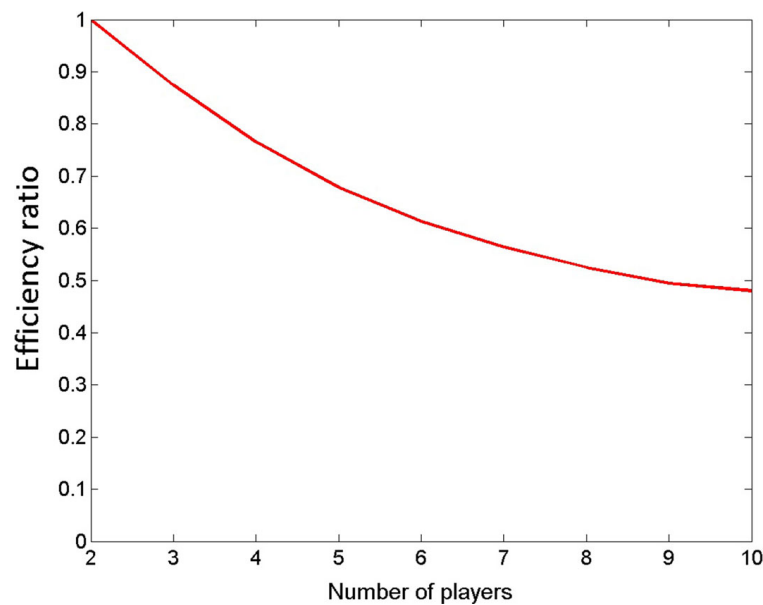
- (b) For all  $n \geq 3$ , the power of Mover 1 ... Mover  $n - 2$  increases monotonically, i.e.,  $\forall_{3 \leq i \leq n-2} \mathbb{P}_{\rho_i}(G) < \mathbb{P}_{\rho_{i+1}}(G)$ .
- (c) For all  $n \geq 2$ , the power of the first mover decreases with  $n$ , i.e.,  $\forall_{n \geq 2} \mathbb{P}_{\rho_1}(G(n)) < \mathbb{P}_{\rho_1}(G(n+1))$ .
- (d) For all  $n \geq 4$ , the power of the last  $n - 1$  movers in game  $G(n)$  increases monotonically w.r.t. the power of  $n - 1$  movers in  $G(n - 1)$ , i.e.,  $\forall_{1 \leq i \leq n-1} \mathbb{P}_{\rho_{n-i}}(G(n-1)) < \mathbb{P}_{\rho_{n-i}}(G(n))$ .
- (e) For all  $n \geq 4$ , the power of an average agent in the game increases with  $n$ , i.e.,  $\forall_{n \geq 4} \frac{\sum_{i=1}^n \mathbb{P}_{\rho_i}(G)}{n} < \frac{\sum_{i=1}^{n+1} \mathbb{P}_{\rho_i}(G)}{n+1}$ .

Observation (e) is especially interesting. Counter-intuitively, the more the participants in the game, the more powerful, on average, they become. Should the pattern in Fig. 1 hold for  $n \geq 10$ , we conjecture that, with  $n$  going to infinity, the power of the first mover will converge to a number close to 0.5, that is, this mover gradually becomes powerless. At the same time, the power of the last two movers converges to a number close to 1, that is, their power becomes close to dictatorship. In addition, since the power of all the remaining movers (from Mover 2 to Mover  $n - 2$ ) increases monotonically, it can be expected that the average power in the system, with  $n$  going to infinity, converges to  $\approx 0.75$ .

This analysis leads to interesting strategic conclusions: if new players are to be added to the game, then, from the point of view of an incumbent player, it is best to add them at the first positions in the order. Furthermore, it is best for the incumbents to enlarge the game as much as possible.

Let us now study the relationship between power and efficiency.

**Efficiency:** In Fig. 2, we plot the average efficiency ratios for games of  $n = 2, \dots, 9$  players. We observe that the efficiency ratio is decreasing with growing  $n$ . The reasons for this can be



**Fig. 2** The average efficiency ratios for games of different size ( $\delta = 0$ )

understood via the distribution of power in the game. As is visible in Fig. 1, with growing  $n$ , the discrepancies between the power of agents increase. This means that more often powerful agents are able to secure favourable outcomes at the expense of powerless agents—a conflict which results in the overall efficiency loss.

*Example 8* Consider the example in Table 2(a). Here, Mover 2 is able to secure his most preferred structure  $\{1\}\{2, 3\}$ , but this structure is ranked 5 and 2 by Mover 1 and Mover 3, respectively. This yields a welfare of  $5 + 1 + 2 = 8$ . Should Mover 2 have less power than Mover 1, the grand coalition might be the outcome with a resulting welfare of  $1 + 2 + 1 = 4$ .

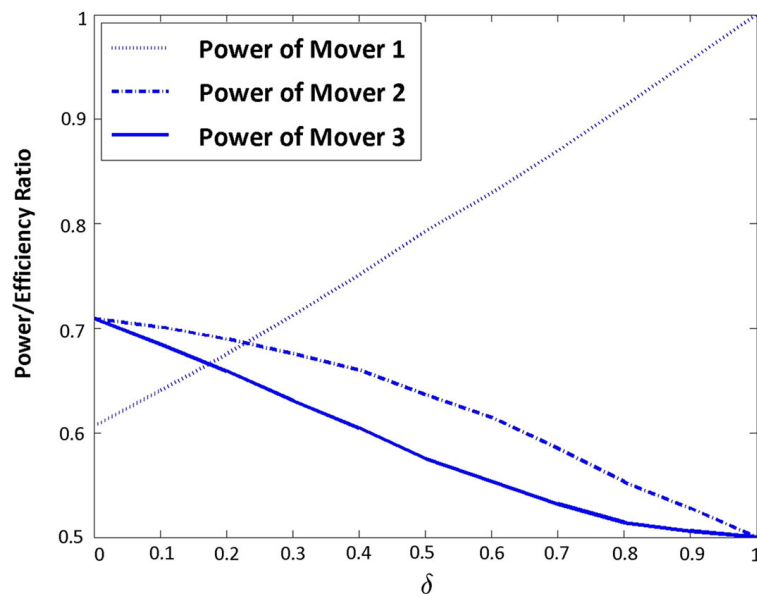
In the next section, we will consider  $\delta \geq 0$  as a potential method to balance power.

#### 4.4 Authoritarianism ( $\delta > 0$ )

In this section, we analyse authoritarian games ( $\delta > 0$ ), including their most extreme case – dictatorship games ( $\delta = 1$ ).

**Power:** Let us consider the 3-player game for which we were able to compute power exactly. In Fig. 3, we plot the power of all the three players for different values of  $\delta$ . As the probability of the random event increases, Mover 1 becomes more and more powerful while Mover 2 and Mover 3 become less powerful. The intuition behind this has already been illustrated in Example 5: for  $\delta = 0.2$  in Table 2(c), Mover 1 knows that Mover 2 opts for the very disliked  $\{1\}\{2, 3\}$ , but he will propose his most preferred grand coalition in the hope that the random event happens (and the grand coalition results). Interpreting the random event from the power perspective, a higher  $\delta$  increases Mover  $i$ 's ( $1 \leq i < n$ ) power to propose, and decreases Mover  $j$ 's ( $i < j \leq n$ ) power to reject.

Analysing Fig. 3 further, for  $\delta \approx 0.2$ , the power of Mover 2 surpasses the powers of the other players. Finally, as expected, for  $\delta = 1$ , Mover 1 becomes the dictator while Mover 2 and Mover 3 become completely powerless. It is important to note that the power of Mover 2



**Fig. 3** Power and efficiency in  $G$  for  $n = 3$

and Mover 3 does not decrease uniformly. For  $\delta > 0$  Mover 2 becomes more powerful than Mover 3. This observation can be formalized as follows:

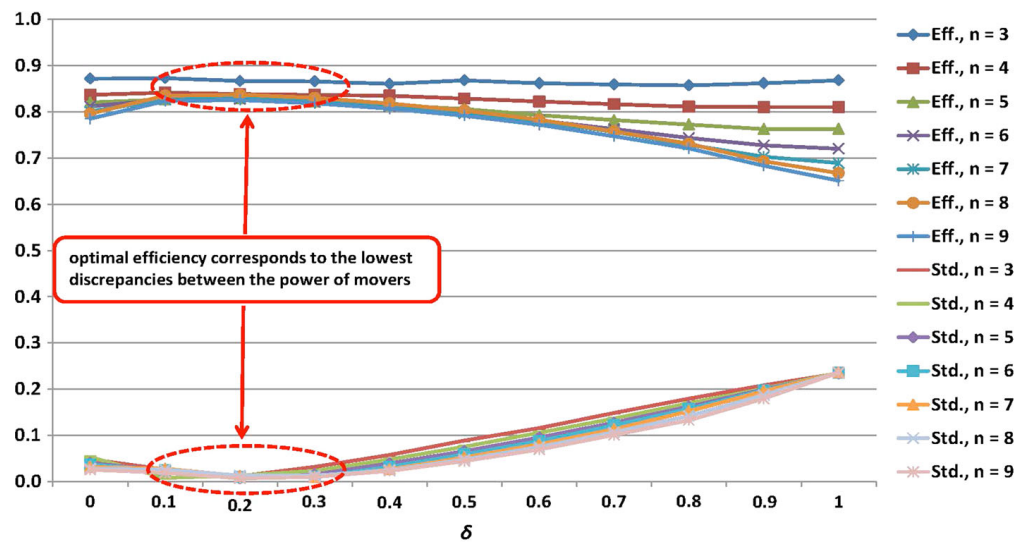
**Proposition 4** For  $\delta > 0$ , the power of Mover  $n - 1$  is always higher than that of Mover  $n$ , i.e.,  $\mathbb{P}_{\rho_{n-1}} > \mathbb{P}_{\rho_n}$ .

*Proof* As we saw in the proof of Proposition 3, if Mover  $n - 1$  prefers not to cooperate with Mover  $n$ , the structure  $\{1\}\{2\} \dots \{n\}$  will be proposed and implemented, irrespective of the preferences of Mover  $n$  and whether or not the random event occurs. While the same holds for Mover  $n$  against Mover  $n - 1$  for  $\delta = 0$ , this is no longer the case for  $\delta > 0$ .

Suppose  $\delta > 0$  and consider the time period  $n - 1$ . In the equilibrium for this time period, all the first  $n - 2$  players on the agenda will be singletons. Hence, let us focus only on how the last two movers structure themselves. Mover  $n - 1$  can propose one of two possible offers:  $\{n - 1, n\}$  or  $\{n - 1\}\{n\}$ . The last mover can only propose  $\{n - 1\}\{n\}$ . With regard to the preferences of the last two movers for these two structures, we can have one of the following four possibilities:

1.  $\{n - 1, n\} \succ_{n-1} \{n - 1\}\{n\}$  and  $\{n - 1, n\} \succ_n \{n - 1\}\{n\}$
2.  $\{n - 1\}\{n\} \succ_{n-1} \{n - 1, n\}$  and  $\{n - 1\}\{n\} \succ_n \{n - 1, n\}$
3.  $\{n - 1\}\{n\} \succ_{n-1} \{n - 1, n\}$  and  $\{n - 1, n\} \succ_n \{n - 1\}\{n\}$
4.  $\{n - 1, n\} \succ_{n-1} \{n - 1\}\{n\}$  and  $\{n - 1\}\{n\} \succ_n \{n - 1, n\}$

For the first two cases, the last two movers have identical power. For the third case, the equilibrium for  $t = n - 1$  will be  $\{n - 1\}\{n\}$  which is better for Mover  $n - 1$  relative to Mover  $n$ . For the fourth case, Mover  $n - 1$  will maximise his expected utility and can propose  $\{n - 1, n\}$  in the hope that the game will end at  $t = n - 1$  due to the random event. In this case, Mover  $n - 1$  will have more power than Mover  $n$ . In other words, whenever Mover  $n - 1$  likes cooperation with Mover  $n$  while Mover  $n$  does not, Mover  $n - 1$  will propose  $\{1\}\{2\} \dots \{n - 1, n\}$  in the hope that the random event happens and this coalition structure gets implemented. Therefore, consider all four cases, the expected rank of Mover  $n - 1$  will be higher than the expected rank of Mover  $n$ .



**Fig. 4** Efficiency versus the standard deviation of power in the game,  $n = 3, 4, \dots, 9$  and  $0 \leq \delta \leq 1$

Note that, for  $\delta > 0$ , we only considered the last two time periods. This is enough because, the equilibrium structures for the first  $n - 1$  time periods will be symmetric in terms of their ranks for Mover  $n - 1$  and Mover  $n$ .  $\square$

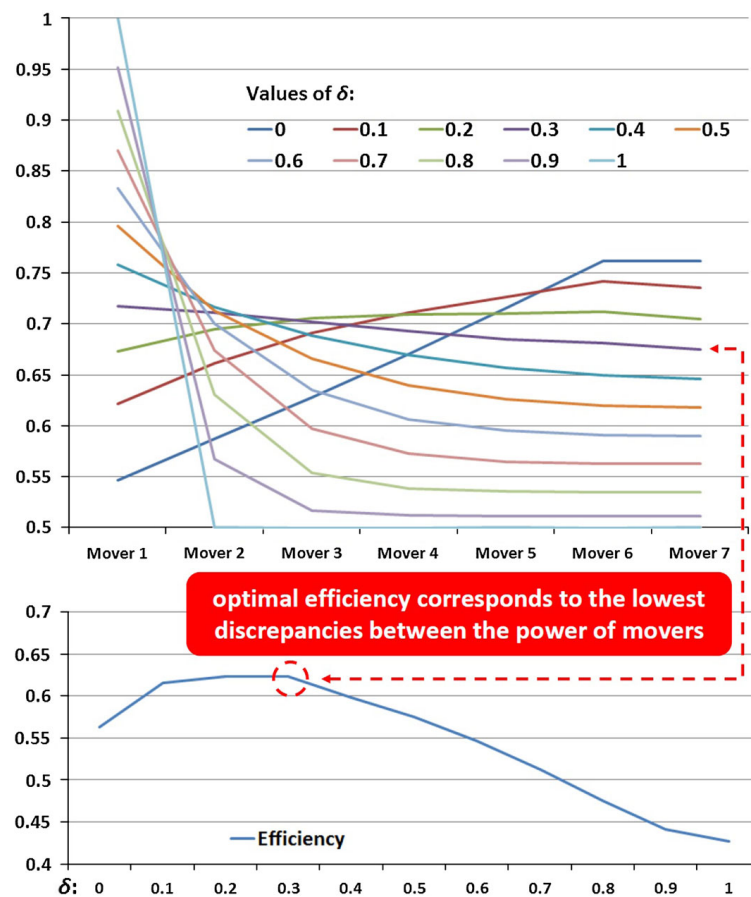
Let us now analyse how increasing the probability of the random event influences the distribution of power among the players. We conducted this analysis for games of up to 9 players and observed the same distribution of power. Below, we will explain the results for  $n = 7$  player games for which the distribution is shown in Fig. 5. Initially, for small  $\delta$  (such as 0.1 and 0.2), the first mover has the lowest power among all and power grows monotonically as we move down the agenda to the subsequent movers. However, the differences between the powers of movers become smaller and smaller. Interestingly, Proposition 4 implies that Mover  $n - 1$  now becomes the most powerful player in the game. As  $\delta$  increases, there comes a turning point when the first mover becomes most powerful. For instance, at  $\delta = 0.4$ , the first mover is the most powerful while Mover  $n$  is the least powerful. Finally, similar to the game of  $n = 3$  players, when  $\delta = 1$  the first mover becomes the dictator and all the other movers are completely powerless.

Next, we will see how  $\delta > 0$  influences efficiency.

**Efficiency:** We suggested in Sect. 4.3 that the decreasing efficiency with increasing  $n$  could be caused by the increasing discrepancies between the powers of the players. The results in Fig. 4 match this conjecture. In particular, the figure compares the efficiency with the standard deviation of players' power for  $n = 3, 4, \dots, 9$ , and  $0 \leq \delta \leq 1$ . Furthermore, in Fig. 5, we zoom to the case  $n = 7$ . We observe that the lowest levels of efficiency are reached for those  $\delta$ 's that correspond to the smallest standard deviations of players' power. In other words, the lower the differences between players' power are, the more desirable the outcome is from a social point of view. Notably, efficiency is the lowest for  $\delta = 1$ .

#### 4.5 Bargaining with a non-cooperative/cooperative player

Until now, while analysing the expected power of an agent, we assumed that all combinations of preferences are equally likely. Such an approach is appropriate from the perspective of a



**Fig. 5** Efficiency versus power in the game, for  $n = 7$  and  $0 \leq \delta \leq 1$

system designer, who does not favour some agents over the others. In this section, we analyse the game from the perspective of a single agent. In particular, we first assume that this agent is *non-cooperative* and dislikes cooperating with other agents down the agenda. Next, we will assume that this player is *cooperative*, that is, he prefers cooperation to non-cooperation. In both cases, this distinguished player will be indexed with 1. Player 1 does not know the preferences of the other agents.

For games with a single non-cooperative/cooperative player, we want to find how the position of this player on the agenda affects his own power, the power of the other players, and the social welfare. To this end, let us consider a game of  $n = 6$  agents and suppose that player 1 is the first mover.

**Player 1 is non-cooperative:** The preferences of the non-cooperative player 1 are fixed as follows. Let us consider any two coalition structures:

$$\pi_1 = \{C_1^1, C_2^1, \dots, C_{|\pi_1|}^1\}, \pi_2 = \{C_1^2, C_2^2, \dots, C_{|\pi_2|}^2\} \in \Pi(N).$$

Without loss of generality, we can assume that the coalitions in  $\pi_1$  and in  $\pi_2$  are ordered from the biggest to the smallest in size and then, if two coalitions have the same size, they are ordered lexicographically. Then, the preference ordering obeys the following rules:

1. If the number of coalitions in  $\pi_1$  is greater than the number of coalitions in  $\pi_2$ , i.e.  $|\pi_1| > |\pi_2|$  then non-cooperative Player 1 prefers  $\pi_1$  to  $\pi_2$ , and *vice versa*.

2. If the number of coalitions in  $\pi_1$  is equal to the number of coalitions in  $\pi_2$  then we compare the vector of sizes of coalitions in  $\pi_1$  with the vector of sizes of coalitions in  $\pi_2$ , and choose the one that is lexicographically less. For instance, if  $\pi_1 = \{1, 2, 3\}\{4, 5\}\{6, 7\}$  and  $\pi_2 = \{1, 2, 3\}\{4, 5, 6\}\{7\}$  then the non-cooperative player 1 prefers  $\pi_1$  to  $\pi_2$ .
3. If the vectors of sizes of coalitions in  $\pi_1$  and  $\pi_2$  are the same, we compare lexicographically the members of  $C_1^1$  to  $C_1^2$  and we choose the coalition structure in which the first coalition is smaller. If  $C_1^1$  and  $C_1^2$  are equal, we compare  $C_2^1$  and  $C_2^2$  and choose the smaller, and so on.

**Player 1 is cooperative:** Here, the preferences are the reverse of the preferences when player 1 is non-cooperative.

In other words, in our 6-player setting, the non-cooperative player 1 ranks the non-cooperative coalition structure  $\{\{1\}, \dots, \{6\}\}$  first. Next, he likes the coalition structure where only players 5 and 6 cooperate, then where 4 and 6, and so on. He likes grand coalition the least. The preferences of the cooperative player 1 are exactly the opposite. Finally, for the other agents, we assume that all combinations of preferences may occur with equal probability.

To focus attention, we will consider two specific agendas: one in which the non-cooperative/cooperative player is placed first, and the other where he is placed last, i.e., the agendas  $(1, 2, 3, 4, 5, 6)$  and  $(2, 3, 4, 5, 6, 1)$ .<sup>5</sup> For these two agendas, we will investigate the power attained by each of the players on the agenda. Looking from the perspective of the non-cooperative/cooperative player, we will seek the most profitable positions for him. Conversely, from the perspective of the society as a whole, we will see where the non-cooperative/cooperative agent should be placed to secure optimal social welfare.

#### 4.5.1 The non-cooperative player 1

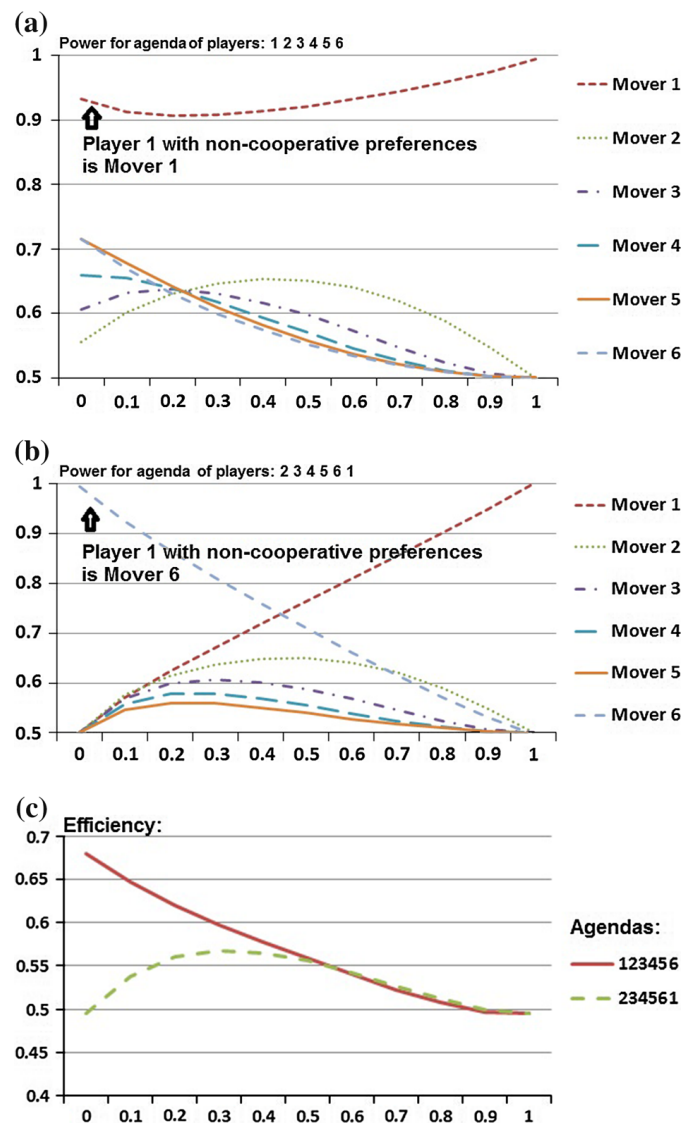
For the two possible agendas, Fig. 6 shows the powers of all the six players and the efficiency ratios for various values of  $\delta$ . In democracy, ( $\delta = 0$ ), player 1 has higher power as the last mover than he has as the first mover. However, from  $\delta = 0.2$  onwards, it is best for player 1 to be the first mover. From a social perspective, it is better to make player 1 the first mover rather than the last mover. The intuition behind this result is as follows. For  $\delta = 0$ , player 1 as the last mover is able to reject any proposal and secure the non-cooperative coalition structure as the outcome of bargaining. As a result, all other players are completely powerless and such a situation is clearly socially inefficient. On the other hand, placing player 1 first actually puts his preferences in line with the working of the game. This is because, if player 1 is rejected and “forced” not to cooperate, this is what this player actually prefers. Thus, agenda  $(1, 2, 3, 4, 5, 6)$  reduces the power of player 1 only a bit, but allows the other players to secure much better coalition structures than for the agenda  $(1, 2, 3, 4, 5, 6)$ .

Figure 6(c), additionally to efficiency ratios for  $(1, 2, 3, 4, 5, 6)$  and  $(2, 3, 4, 5, 6, 1)$ , present the ratios for the remaining agendas, i.e.,  $(2, 1, 3, 4, 5, 6)$ ,  $(2, 3, 1, 4, 5, 6)$ ,  $(2, 3, 4, 1, 5, 6)$ , and  $(2, 3, 4, 5, 1, 6)$ . We observe a gradual decrease of social efficiency with the decreasing position of the non-cooperative player in the agenda.

Finally, we note that qualitatively the same results were obtained also for other numbers of players in the game:  $n = 5, 7, 8$ , and 9.

<sup>5</sup> Since all players other than 1 are symmetric,  $(1, 2, 3, 4, 5, 6)$  represents all possible agendas, where the non-cooperative/cooperative player is the first, and  $(2, 3, 4, 5, 6, 1)$ , where he is the last.



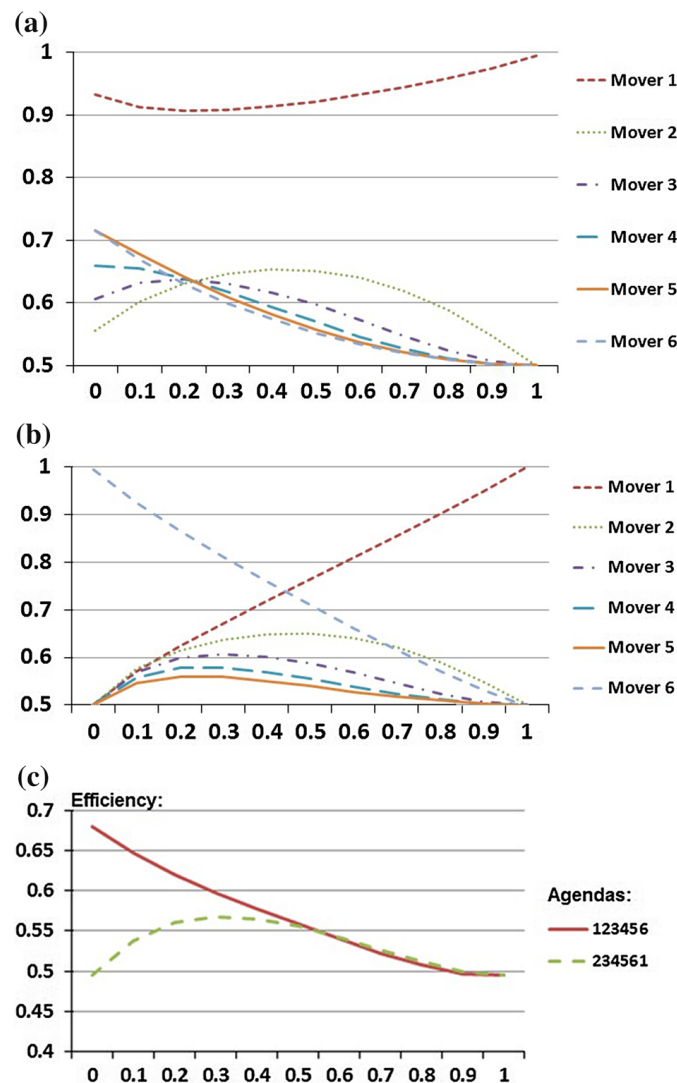


**Fig. 6** Power and efficiency for  $n = 6$  and agendas (1, 2, 3, 4, 5, 6) and (2, 3, 4, 5, 6, 1). Player 1 is *a priori* non-cooperative. **a** Power for the agenda (1, 2, 3, 4, 5, 6). **b** Power for the agenda (2, 3, 4, 5, 6, 1). **c** Efficiency

#### 4.5.2 The cooperative player 1

Figure 7 shows the corresponding results for the cooperative player 1. Now, in democracy ( $\delta = 0$ ), player 1 has lower power as the last mover than he has as the first mover. Clearly, as the last mover, player 1 has no chance to exercise his power to propose while his power to reject is also significantly hampered. Also, we note that in both cases, that is, when player 1 is the first or the last, his power is always substantially lower than the powers of all the other players. In fact, when he is last, for any  $\delta$  his power is always lower than 0.5 and the lowest among all the players.

From the social perspective, the shape of the efficiency diagram in Fig. 7 is similar to that in Fig. 6, but for each data point, the social welfare is higher when player 1 is cooperative compared to when he is non-cooperative. In other words, it is always better for other players



**Fig. 7** Power and efficiency for  $n = 6$  and agendas (1, 2, 3, 4, 5, 6) and (2, 3, 4, 5, 6, 1). Player 1 is *a priori* cooperative. **a** Power for the agenda (1, 2, 3, 4, 5, 6). Here, **Player 1 = Mover 1** (b). Power for the agenda (2, 3, 4, 5, 6, 1) (c) Efficiency

to deal with the cooperative player than with the non-cooperative one. Furthermore, for values of  $\delta \leq 0.2$  agenda (1, 2, 3, 4, 5, 6) is the most efficient while for  $\delta > 0.2$  agenda (1, 2, 3, 4, 5, 6). The agendas intermediate between these two have typically lower social efficiency, with the exception of (2, 1, 3, 4, 5, 6) for  $\delta \geq 0.9$ .

Finally, we note that also in this case qualitatively the same results were obtained also for other numbers of players in the game:  $n = 5, 7, 8$ , and 9.

#### 4.6 Bargaining with binary games

The results obtained thus far are for coalition structure games where each agent's rank is a bijective map from  $\Pi(N)$  to  $\{1, \dots, \text{Bell}(n)\}$ . Our aim now is to determine if these results still hold in a simpler setting, where each  $r_i$  is a map from  $\Pi(N)$  to  $\{0, 1\}$  with  $r_i(\pi) = 1$

( $r_i(\pi) = 0$ ) meaning that player  $i$  likes (dislikes)  $\pi$ . Importantly, player  $i$  *equally* likes all coalition structures ranked  $r_i(\pi) = 1$ , and equally dislikes all the others. We call such games *binary games*, and denote them by  $G^b$ .

Let us first consider a few extreme preference orderings in binary games:

- (a) An agent  $i$  likes all coalition structures. In this case  $\forall_{\pi \in \Pi(N)} r_i(\pi) = 1$ .
- (b) Conversely,  $i$  dislikes all coalition structures. In this case  $\forall_{\pi \in \Pi(N)} r_i(\pi) = 0$ .
- (c) Furthermore,  $i$  is single minded, that is,  $i$  likes exactly one coalition structure and dislikes all the others. That is,  $\exists_{\pi^* \in \Pi(N)} r_i(\pi^*) = 1 \wedge \forall_{\pi \in \Pi(N) \setminus \{\pi^*\}} r_i(\pi) = 0$ .

In general, any player can like an arbitrary subset of coalition structures and dislike all the remaining ones. Thus, the number of all possible combinations of preference orderings is  $(2^{B_n})^n$ . One way to generate a preference ordering for a player  $i$  in a binary game is to assume that a fraction  $0 \leq \omega \leq 1$  of all coalition structure is liked by  $i$  and a fraction  $1 - \omega$  is dislikes. In case (a) above we have  $\omega = 1$ , in (b)  $\omega = 0$ . On the other hand, in (c) we have  $\omega = \frac{1}{B_n}$ .

To focus attention, in what follows we will assume that  $\omega = \frac{1}{2}$ , that is, any player likes a half of the coalition structures and dislikes the remaining half. This also has a probabilistic interpretation that any coalition structure has equal probability of being liked or disliked. If the number of coalition structures is odd, as it happens, for instance, for  $n = 3$ , we assume that any player with certainty likes  $\lfloor \frac{1}{2} B_n \rfloor$  coalition structures and dislikes exactly the same number of them. A player likes the remaining one coalition structure (note that  $B_n - 2 \times \lfloor \frac{1}{2} B_n \rfloor$ ) with probability  $\frac{1}{2}$ .

Let us consider one possible way of converting preferences from the standard bargaining game considered previously to the corresponding preferences of a binary game. Given, for example,  $n = 3$  and  $B_3 = 5$ , we can assume that either the first 2 or the first 3 coalition structures ranked top in a standard bargaining game are liked by any player in a binary game while the remaining coalition structures are disliked.

Using the above method, let us convert the game in Example 9 into a binary game.

**Example 9** The players' rankings for all possible coalition structures in Table 2(a) are converted to binary rankings. In particular, we assume that, in the binary game, Mover 1 likes those coalition structures that are his top 3 in the standard bargaining game, and dislikes the remaining structures. Likewise for Mover 3. On the other hand, Mover 2 likes only his top two ranked coalition structures and dislikes the remaining ones.

At first sight, it may seem that if for an arbitrary  $0 \leq \omega \leq 1$  we convert a standard bargaining game into a binary bargaining game then both games are strategically equivalent. For instance, the binary game in Example 9 may seem to be strategically equivalent to the game from Table 2(a). After all, in a standard bargaining game the players strive to achieve the best possible ranking. Since only preferences change but not rules of the bargaining game, the players are able to achieve the same coalition structures. Thus, for a given  $\omega$  in a binary game, the players are either able to achieve the coalition structure that they like or not.

Interestingly, the above intuition is flawed, i.e., a standard bargaining game is not strategically equivalent to the corresponding binary game (for a given  $\omega$ ). This is due to the fact that the players become indifferent with respect to which coalition structure they like is created. Therefore, certain rejections that occur in a standard bargaining game do not occur in a binary one. Let us consider again the games from Example 9. We know that, for the standard game (see Table 4), Mover 1's equilibrium offer will be the structure  $\{1\}\{2, 3\}$  because Mover 2 prefers it the most and he will reject any other offer made by Mover 1. Next, consider the

**Table 4** The players' rankings for all possible coalition structures from Table 2(a) converted into the rankings of a binary game

Game Coalition structure	Table 2(a)			Table 2(a) as binary game		
	Mover 1	Mover 2	Mover 3	Mover 1	Mover 2	Mover 3
{1, 2, 3}	1	2	1	1	1	1
{1, 2}{3}	4	4	4	0	0	0
{1, 3}{2}	3	3	3	1	0	1
{1}{2, 3}	5	1	2	0	1	1
{1}{2}{3}	2	5	5	1	0	0

The rankings from Table 2(a) are reproduced for easy comparison

**Table 5** An illustration of preferences that lead to multiple equilibria in a binary game

Coalition structure	A			B		
	Mover 1	Mover 2	Mover 3	Mover 1	Mover 2	Mover 3
{1, 2, 3}	1	1	0	0	1	0
{1, 2}{3}	1	1	0	1	1	0
{1, 3}{2}	1	0	1	1	1	1
{1}{2, 3}	0	1	0	0	1	0
{1}{2}{3}	1	0	0	1	0	0

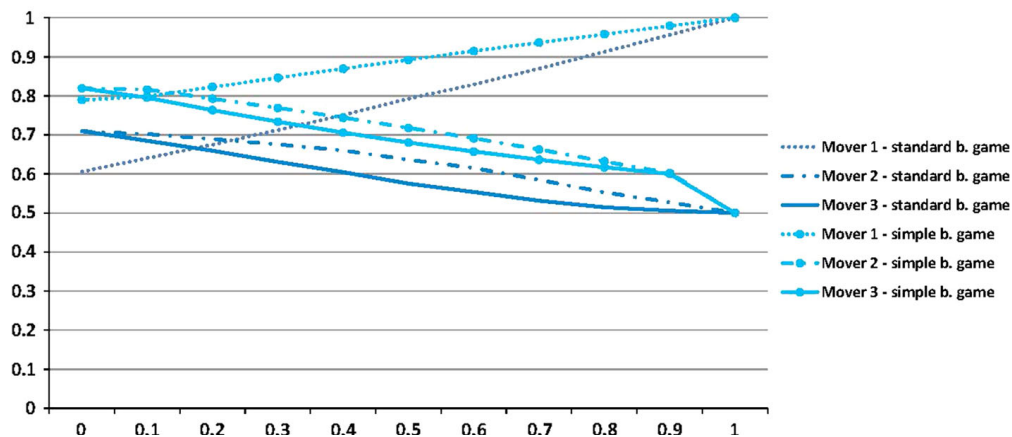
preferences given in Table 4 for the corresponding binary game. Here, Mover 1's equilibrium offer will be the structure {1, 2, 3} because Mover 2 likes {1}{2, 3} and {1, 2, 3} equally and is indifferent between them.

A noteworthy feature of binary games is that, unlike standard games, they can give rise to *multiple equilibrium offers*. This is illustrated by means of the preferences listed in Table 5. Suppose  $\delta = 0$ . Consider the preferences under the columns labelled A. In the second time period, Mover 2 is able to secure a coalition he likes. Given this, Mover 1 can either propose {1, 2, 3} or {1, 2}{3}. In order to resolve such ties, we use the following tie breaking rule: ties are broken in favor of structures that maximize social welfare. After applying this rule, ties that still persist are broken in favor of those structures that can be offered only by the proposer but not by those that succeed him on the agenda (recall that Mover 1 can offer any coalition structure, Mover 2 can only offer those structures in which Mover 1 is a singleton, and Mover 3 can only offer the structure with three singletons). After applying this rule, ties that still persist are broken randomly as the choice does not affect player preferences. As per this rule, Mover 1 will make a random choice between {1, 2, 3} and {1, 2}{3} because both structures are equally good in terms of welfare and only Mover 1 can offer both structures. Next, consider the preferences under the columns labelled B. Here, Mover 1 will break the tie between {1, 2}{3} and {1, 3}{2} in favor of the latter because it is better in terms of welfare.

In order to analyse power and welfare in binary games we redefine the power index and the efficiency ratio as follows:

**Definition 6** Mover  $i$ 's ( $i \in N$ ) power index over the set of binary games in  $\mathcal{G}^s$  is defined as:

$$\mathbb{P}_{\rho_i}(\mathcal{G}^s) = \mathbb{E}_{\rho_i}^s(\mathcal{G}^s),$$



**Fig. 8** The power in a three-player binary game compared to the power in a three-player standard bargaining game

where  $\mathbb{E}_{\rho_i}^s(\mathcal{G}^s)$  denotes Mover  $i$ 's average expected rank in the equilibrium for the binary games in  $\mathcal{G}^s$  and it is defined as follows:

$$\mathbb{E}_{\rho_i}^s(\mathcal{G}^s) = \frac{1}{(\text{Bell}(n)!)^n} \sum_{G^b \in \mathcal{G}^s} er_{\rho_i}(\pi_1^*(G^b)).$$

Note that the power index for binary games also takes values between 0 and 1, i.e.,  $0 \leq \mathbb{P}_{\rho_i}(\mathcal{G}^s) \leq 1$ . The closer a player's power index to 1, the more powerful he is. Intuitively, Mover  $i$  is able to secure a coalition structure that he likes more often.

As far as the efficiency ratio is concerned, we have:

**Definition 7** For a given  $n$  and  $\delta$ , the efficiency ratio  $\mathcal{E}(G^b)$  for a binary game  $G^b$  is the ratio of the sum of the players' ranks for the bargained structure in  $G^b$  and the sum of ranks for the globally optimal coalition structure, i.e., we have:

$$\mathcal{E}(G^b) = \left( \sum_{i=1}^n er_i(\pi_1^*(G^b)) \right) / \left( \sum_{i=1}^n r_i(\pi_{SW}(G^b)) \right).$$

Although the standard and binary games are not strategically equivalent, in our setting for  $\omega = \frac{1}{2}$ , they are characterised by a very similar distribution of power between movers. Figure 8 compares the power in a three-player binary game to the power in a three-player standard bargaining game. The results are similar. The results concerning efficiency are also very similar.

Before closing this section, we will comment on the generality of our results. We considered two types of rankings: one where each coalition structure has a unique rank, and the other where each structure has one of two possible ranks. For the former, an agent's expected rank is the average rank over all games played. Those situations where it is not possible to take simple averages need further investigation. For example, consider a three player game for which five coalition structures are possible and therefore five possible ranks. Suppose the game is played two times. Suppose also that the outcome of the first play is a structure that a player prefers the most, and the outcome of the second play results in his least preferred structure. Then his average rank will be  $(1+5)/2=3$ . This is the same as getting the rank 2 in the first play and rank 4 in the second because the average rank will still be 3. We assumed

that these two situations the same because the average rank is the same. However, in general, it is possible for a player to prefer the first situation (i.e., getting rank 1 in the first play and rank 5 in the second) more/ less than the second situation (i.e., getting rank 2 in the first play and rank 4 in the second). Such situations need further investigation. It is worth noting that, such a problem does not arise for binary preferences. Since there are only two possible ranks, an agent's expected rank depends on the number of resulting outcomes that a player likes.

## 5 Non-cooperative equilibrium and the core

Our attention so far was focussed on the analysis of non-cooperative bargaining equilibrium. However, it is important to understand how this non-cooperative equilibrium relates to the cooperative equilibria<sup>6</sup> for our coalition structure game. In this article, we will focus on one of the key solution concepts from cooperative game theory called the *core* and relate it to the non-cooperative bargaining equilibrium. The definition of core for coalitional games is unique but for coalition structure games with externalities, a number of definitions of the core have been proposed [21]. The earliest and a very widely studied notion is that of  $\alpha$  core [3]. In the context of the coalition structure games defined in Sect. 2.1 an  $\alpha$  core (or  $\alpha$  stable) structure is defined as follows:

**Definition 8** A coalition structure  $\pi$  is  $\alpha$  stable if there does not exist a coalition  $C \subseteq N$  and a partition  $\pi'$  of  $C$  such that,  $\forall i \in C, \forall \pi_{N \setminus C} \in \Pi_{N \setminus C}, r_i(\pi'_C \cup \pi_{N \setminus C}) < r_i(\pi)$ .

In words, a coalition structure is  $\alpha$  stable if there is no coalition  $C$  that has a partition that is more preferred by all the agents in  $C$  regardless of how the players outside of  $C$  are partitioned. Theorem 6 is a characterization of the relation between the non-cooperative equilibrium for the bargaining game defined in Algorithm 1 and the  $\alpha$  stable structure of the CSG defined in Sect. 2.1.

**Theorem 2** *The bargained outcome of the noncooperative game may or may not be  $\alpha$  stable.*

*Proof* The proof is by example: one for the case where the noncooperative equilibrium is  $\alpha$  stable and the other for the case where it is not. Consider the preferences given in Table 6A and suppose that  $\delta = 0$ . In the noncooperative equilibrium for the agenda 123, the players will form the structure  $\pi = \{1\}\{2, 3\}$ . This structure is most preferred by Mover 2 and Mover 3. Hence, they will neither like to break apart nor allow Mover 1 into their coalition. Thus,  $\pi = \{1\}, \{2, 3\}$  is  $\alpha$  stable since there does not exist a coalition  $C$  and a partition  $\pi'$  of  $C$  such that,  $\forall i \in C, \forall \pi_{N \setminus C} \in \Pi_{N \setminus C}, r_i(\pi'_C \cup \pi_{N \setminus C}) < r_i(\pi)$ .

Now consider the preferences given in Table 6B and suppose that  $\delta = 0$ . In the noncooperative equilibrium for the agenda 123, the players will form the structure  $\pi = \{1\}\{2\}\{3\}$ . This structure is not  $\alpha$  stable since for the coalition  $C = \{1, 2\}$  and the partition  $\pi' = \{1, 2\}$  of  $C$ , the condition  $\forall i \in C, \forall \pi_{N \setminus C} \in \Pi_{N \setminus C}, r_i(\pi'_C \cup \pi_{N \setminus C}) < r_i(\pi)$  is false. In words, compared to the structure  $\{1\}\{2\}\{3\}$ , the structure  $\{1, 2\}\{3\}$  is better for both player 1 and for player 2. Hence  $\pi = \{1\}\{2\}\{3\}$  is not  $\alpha$  stable.  $\square$

Theorem 2 raises the following questions:

1. Under what conditions is the noncooperative equilibrium  $\alpha$  stable?
2. For any coalition structure game, does an agenda always exist such that the noncooperative equilibrium coincides with the  $\alpha$  core?

Further research is needed to answer these questions.

<sup>6</sup> We thank the anonymous reviewer for suggesting the need for this section.



**Table 6** A: An illustration of preferences leading to a bargaining equilibrium (shown in boldface) that is  $\alpha$  stable

Coalition structure	A			B		
	Mover 1	Mover 2	Mover 3	Mover 1	Mover 2	Mover 3
{1, 2, 3}	1	2	5	5	2	1
{1, 2}{3}	2	5	4	1	3	4
{1, 3}{2}	5	3	3	3	1	5
{1}{2, 3}	<b>3</b>	<b>1</b>	<b>1</b>	4	5	2
{1}{2}{3}	4	4	2	<b>2</b>	<b>4</b>	<b>3</b>

B: an illustration of preferences leading to a bargaining equilibrium (shown in boldface) that is not  $\alpha$  stable

## 6 Related work

There is a large body of work focused on finding stable coalitions (as opposed to our focus on the formation of a coalition structure) through negotiation. This includes Shehory and Kraus [41] who studied coalition formation through a process of negotiation. In each stage of the negotiation, the agents propose coalitions and not coalition structures. The authors show that the agreement that the agents reach is a modified version of the kernel. Later, Kraus et al. [29] studied coalition formation in the context of task allocation in multiagent systems. They considered uncertainty about coalition values and analyzed a bargaining protocol in which the agents propose coalitions (contrast this with our bargaining game in which the players propose coalition structures). In order to handle the complexity of computing optimal offers, they proposed heuristic strategies. By assuming that the agents are truthful, they showed the stability of the heuristic strategies through simulations. Blankenburg et al. [6] also considered uncertainty about coalition values and combined the notion of kernel from game theory with fuzzy sets to define a new concept called fuzzy kernel. They provided a bilateral coalition formation algorithm for forming fuzzy kernel stable coalitions among agents.

Bloch and Rottier [9] considered coalition formation in the context of games where the players are political parties that bargain noncooperatively to form a government. The authors focused on transferable utility games without externalities and analyzed the impact of the agenda on the outcome of bargaining. Here, the players bargain to form one winning coalition. In contrast, we consider non transferable utility games with externalities. In addition, our noncooperative game allows the players to bargain a *coalition structure*, not just a winning coalition.

Hart and Mas-Colell [22] also studied strategic bargaining for the formation of a single coalition in the context of a protocol that proceeds in rounds and, in each round, the proposing agent is selected at random. The authors showed the relationship between the outcome of this game and the Shapley value.

Recently, Fatima and Woodridge [15] investigated majority bargaining for resource division. In this model, there is uncertainty about the players' weights and the agenda is defined exogenously. They showed that the non-cooperative subgame perfect equilibrium results in an instant, unique, and Pareto optimal agreement. Their work uncovered a number of important properties of the non-cooperative equilibrium such as its computational complexity and how this complexity is effected by changes to the conflict point. They also exposed the conditions for the non-cooperative equilibrium to be core stable and for it to be nucleolus stable. In addition, they showed how an individual's bargained share of the resource is influenced by his position on the agenda.

In contrast to the aforementioned literature, we are interested in generating a coalition structure and not just a coalition for which the existing work can be classified as per Table 1. In the context of this table, a number of non-strategic models are aimed at generating a coalition structure that is optimal at a system-wide level (for a thorough analysis of the computational complexity of the optimal coalition structure generation problem, see [4]). For a survey of these models, see [34]. One key assumption that underlies all the non-strategic models is that a central entity is able to impose a socially optimal structure on all the agents. Their focus is only on generating an optimal structure regardless of whether or not it will be acceptable to the individual agents. In other words, no consideration is given to whether such a structure is stable or not. Although such an assumption simplifies the task of generating an optimal structure, it does not hold for multiagent systems. In a multiagent system, there is no central entity empowered to impose his decision on other agents. Each individual makes his own decision and has his own preferences that often conflict with others'. Consequently, it is important to study how conflicts can be resolved and stable structures formed through a strategic process of bargaining. Our research is aimed at achieving this goal.

Within the class of strategic models, [7, 8] took a non-cooperative approach while [10, 12, 36] used a cooperative approach. A key cooperative notion of stability is the *kernel*. Klusch and Shehory [26] devised an polynomial anytime algorithm for generating kernel stable coalitions. However, this work does not deal with the generation of stable coalition structures in the presence of externalities.

Bogomolnaia and Jackson [10] considered the problem of partitioning a society into coalitions in the context of purely hedonic games. On the other hand, Diamantoudi and Xue [12] considered games that are a special class of partition function games [42] and studied coalition structure generation using a cooperative approach. Their work is in the spirit of Ray and Vohra's [36] equilibrium binding agreements and von Neumann and Morgenstern's [31] stable set.

Research on the analysis of non-cooperative bargaining for coalition structure formation started relatively recently. A number of authors modeled coalition structure formation as a non-cooperative sequential game. Early, and perhaps the most well-known work in this category is by Bloch [7]. In this model with externalities, the first mover (defined by an *exogenous* agenda) starts by proposing a coalition. If all prospective members accept, the coalition is formed. If a player rejects, she becomes the initiator in the next round. An important feature of the game is that, once a coalition is formed, the game is only played among the remaining players and established coalitions may not attract new members nor break apart. Hence, by agreeing to group in a coalition, players commit to stay in that coalition. Bloch showed that any core stable structure can be attained as the stationary equilibrium of the noncooperative game if the latter is non-empty.

Bloch and Diamantoudi [8] analysed a bargaining protocol for a class of non-transferable utility games called hedonic games. In a hedonic game, there is a preference ordering, one for every agent, over every coalition. An agent wants to join a coalition for the pleasure of the company of its members, i.e., there are no externalities. The bargaining is conducted as per a sequential alternating offers game with an exogenous agenda. The proposer makes an offer. An offer can be a coalition  $S$  to which the proposer belongs or another player who is then recognised to make an offer. In other words, the game offers the players the possibility to pass the buck to another player. If a coalition  $S$  is proposed, all its members respond sequentially to the offer. If all of them accept, the coalition  $S$  leaves and the game continues with just the remaining players. The first player in the remaining ones makes a proposal, etc. If one of the players rejects the offer, the coalition is not formed and the set of active players remains the same. The rejector then becomes the proposer in the next time period. The authors showed

that under certain conditions, this protocol leads to stationary equilibrium outcomes that belong to the core of the hedonic game.

For coalitional resource games (a form of non-transferable utility games), Dunne et al. [13] analyzed a protocol for bargaining cooperation structures. Unlike our protocol, the number of bargaining rounds in theirs is exactly equal to the number of bargainers, so there is zero probability that the game ends in a given time period other than the last one. They showed that the negotiation outcome satisfies several desirable properties: Pareto optimality, dummy player, and pseudo-symmetry. In contrast, our focus is on the relation between a players position on the agenda and his bargaining power. We also analyse the relationship between the distribution of the power of the individual players, the level of democracy, and the welfare efficiency of the protocol.

In the noncooperative models discussed above, the bargained outcome may not be as efficient as the *social optimum*. In the context of networks, Korilis et al [27] gave the relation between the noncooperative Nash equilibrium and the global optimum. In order to measure how far a bargained outcome will be from a social optimum, Koutsoupias and Papadimitriou [28] defined the term *price of anarchy* and measured it in terms of the ratio of the worst possible Nash equilibrium and a social optimum. They derived upper and lower bounds for this ratio for resource sharing in network games. While they used the ratio of the worst possible Nash equilibrium and the social optimum as a measure of the price of anarchy, we use the ratio of the average equilibrium and the social optimum. In addition, they derived theoretical bounds but we provide analysis through simulations. Lastly, while they dealt with network games, we deal with coalition structure games.

Our focus has mainly been on the noncooperative equilibrium although we also showed that it can, sometimes, coincide with the cooperative equilibrium notion of  $\alpha$  core. A key feature that distinguishes our research from other noncooperative approaches is that we analyze the relation between power and agenda while none of the above did. The only exception is the paper by Bloch and Rottier [9]. However, their bargaining game differs from ours; the players in their game propose coalitions. In contrast, in our bargaining game, the players propose coalition structures. In addition, Bloch and Rottier considered transferable utility games without externalities while we analysed nontransferable utility games with externalities.

## 7 Conclusions

In this article, we explored a noncooperative game for bargaining a coalition structure. The players have different, possibly conflicting, preferences for the possible coalition structures. This finite-horizon alternating offers game goes through at most  $n$  ( $n$  is the number of players in the game) rounds. In each round, a player is called to propose an offer. The order in which the players are called is given by an exogenously defined agenda. In each time period, the game ends with probability  $\delta \geq 0$ . We first obtained the subgame perfect Nash equilibrium for the game. Then we showed how a player's position on the agenda affects his power. We also analysed the relationship between the distribution of the power of individual players, the level of democracy, and the welfare efficiency of the protocol. Surprisingly, we found that purely democratic games are welfare inefficient and this is due to the uneven distribution of the power of individual players. Interestingly, introducing a degree of authoritarianism into the protocol makes the distribution of power more equitable and also maximizes welfare. These results were found to be invariant to two types of games: those with totally ordered player preferences and those with binary preferences.

Due to the computational complexity of determining the noncooperative equilibrium, our results are limited to small games. One way to address this complexity is to develop approximation methods. Nevertheless, in order to evaluate such approximations, it is crucial to know the exact equilibrium solutions. Our result therefore provide a benchmarking standard. Our results are also useful to analyze bigger games where an individual player represents a group of agents.

Although we considered externalities, we focused on the formation of non overlapping coalitions. Thus, a limitation of our model is that it cannot be applied to situations where a player can be a member of multiple coalitions. We also focused on the bargaining game with an exogenous agenda. However, given that players' powers differ, self interested players might be interested in extending the bargaining mechanism so that the agenda is endogenized.

There are many possible directions for future work. First, we showed that the noncooperative equilibrium structure is only sometimes  $\alpha$  stable. A more extensive analysis is needed to gain a full understanding of the conditions when the noncooperative and the cooperative equilibria coincide. Second, we analyzed one specific bargaining game. There are many ways of altering the rules of this game. For instance, we took the agenda as given. It will be interesting to analyze what happens if the proposer for a given round is chosen randomly. Also, in the proposed bargaining game, we can have a maximum of  $n$  rounds. What happens if we increase this limit remains to be explored. Another extension concerns introducing elements of dynamic coalition formation in the spirit of Arnold and Schwalbe [2] or Klusch and Gerber [25]. Regarding the modeling of player preferences, we considered two possibilities: one where each coalition structure has a unique rank and another where the rank is a binary number. It would be interesting to explore more general player types of preferences. Yet another possibility is to develop approximation methods to overcome the computational complexity of finding the noncooperative bargaining equilibrium for large games.

**Acknowledgments** We thank the reviewers for their detailed and helpful comments. Tomasz Michalak and Michael Wooldridge were supported by the European Research Council under Advanced Grant 291528 ("RACE").

## Appendix: The Monte Carlo sampling procedure

The basic idea of Monte Carlo sampling is to consider only a subset of all games  $\mathcal{G}' \subset \mathcal{G}$  and use  $\mathcal{P}_{\rho_i}(\mathcal{G}')$  and  $\mathbb{E}_{\rho_i}(\mathcal{G}')$  as estimators of  $\mathcal{P}_{\rho_i}(\mathcal{G})$  and  $\mathbb{E}_{\rho_i}(\mathcal{G})$ . Our sampling procedure can be described as follows. For a given  $N, \rho, \delta$ , we first chose the number of Monte Carlo iterations, that is,  $|\mathcal{G}'|$ . Next, for each iteration:

- (i) We sample a game  $G$  from  $\mathcal{G}$  (to be included in  $\mathcal{G}'$ ) by sampling a random combination of preference orderings of all the players in  $N$  (see, for instance, Table 2). To this end use Knuth's shuffle algorithm [14].
- (ii) For the combination of preferences generated in (i), we identify the welfare maximizing coalition structure(s),  $\pi_{SW}(G)$ ;
- (iii) We solve the game with backward induction and compute the equilibrium coalition structure,  $\pi_1^*(G)$ ;
- (iv) For each player  $i \in N$  we record his expected rank for the equilibrium coalition structure,  $er_{\rho_i}(\pi_1^*(G))$ . We also compute the efficiency of the game  $G$ ,  $er_{\rho_i}(\pi_1^*(G))$ .

Finally, having generated  $G$ , we compute both  $\mathcal{P}_{\rho_i}(\mathcal{G}')$  and  $\mathbb{E}_{\rho_i}(\mathcal{G}')$ .

## References

1. Allingham, M. (1975). Economic power and values of games. *Journal of Economics*, 35, 293–299.
2. Arnold, T., & Schwalbe, U. (2002). Dynamic coalition formation and the core. *Journal of Economic Behavior and Organization*, 49(3), 363–380.
3. Auman, R. (1961). The core of a cooperative game without side payments. *Transactions of the American Mathematical Society*, 98(3), 539–552.
4. Aziz, H., & de Keijzer. (2011). Complexity of coalition structure generation. In: *Proceedings of the 10th International Joint Conference on Autonomous Agent and MultiAgent Systems*, (pp. 191–198).
5. Barry, B. (1980). Is it better to be powerful or lucky?: Part I and Part II. *Political Studies*, 28, 183–194. and 338–352.
6. Blankenburg, B., Klusch, M., & Shehory, O. (2003). Fuzzy kernel-stable coalitions between rational agents. In: *Proceedings of AAMAS*, (pp. 9–16).
7. Bloch, F. (1996). Sequential formation of coalitions in games with externalities and fixed payoff division. *Games and Economic Behavior*, 14, 90–123.
8. Bloch, F., & Diamantoudi, E. (2011). Noncooperative formation of coalitions in hedonic games. *International Journal of Game Theory*, 40, 263–280.
9. Bloch, F., & Rottier, S. (2002). Agenda control in coalition formation. *Social Choice Welfare*, 19, 769–788.
10. Bogomolnaia, A., & Jackson, M. O. (2002). The stability of hedonic coalition structures. *Games and Economic Behavior*, 38, 201–230.
11. Chalkiadakis, G., Elkind, E., & Wooldridge, M. (2011). *Computational Aspects of Cooperative Game Theory*. Morgan-Claypool.
12. Diamantoudi, E., & Xu, L. (2007). Coalitions, agreements and efficiency. *Journal of Economic Theory*, 136, 105–125.
13. Dunne, P. E., Kraus, S., Manisterski, E., & Wooldridge, M. (2010). Solving coalitional resource games. *AI Journal*, 174, 20–50.
14. Durstenfeld, R. (1964). Algorithm 235: Random permutation. *Communications of the ACM*, 7(7), 420–421.
15. Fatima, S., & Wooldridge, M. (2015). Majority bargaining for resource division. *Journal of Autonomous Agents and Multiagent Systems*. doi:10.1007/s10458-015-9290-8.
16. Felsenthal, D., & Machover, M. (1988). *The measurement of voting power: Theory and practice, problems and paradoxes*. Cheltenham: Edward Elgar.
17. Fink, A. (2006). Supply chain coordination by means of automated negotiations between autonomous agents. In B. Chaib Draa abd & J. Muller (Eds.), *Multiagent based supply chain management* (pp. 351–372). Berlin: Springer.
18. Finus, M. (2003). New developments in coalition theory: An application to the case of global pollution control. In: *Environmental Policy in an International Perspective*, (pp. 19–49). Netherlands: Kluwer Academic Publishers.
19. Finus, M., & Rundshagen, B. (2009). Membership rules and stability of coalition structures in positive externality games. *Social Choice and Welfare*, 32, 389–406.
20. Gallagher, M. (1992). Comparing proportional representation electoral systems: Quotas, thresholds, paradoxes and majorities. *British Journal of Political Science*, 22(04), 469–496.
21. Hart, S., & Curz, M. (1983). Endogenous formation of coalitions. *Econometrica*, 51, 1047–1064.
22. Hart, S., & Mas-Collel, A. (2010). Bargaining and cooperation in strategic form games. *Journal of the European Economic Association*, 8(1), 7–33.
23. Holler, M., & Owen, G. (2010). Why power indices and coalition formation. In M. Holler & G. Owen (Eds.), *Power indices and coalition formation* (pp. 1–13). The Netherlands: Kluwer Academic Publishers.
24. Huq, G. (2010). *Automated negotiation in multi-agent based electronic business: Negotiation in business-to-business (B2B) transactions in supply chain management based electronic business*. Miller: VDM Verlag Dr.
25. Klusch, M., & Gerber, A. (2002). Dynamic coalition formation among rational agents. *IEEE Intelligent Systems*, 17(3), 42–47.
26. Klusch, M., & Shehory, O. (1996). A polynomial kernel-oriented coalition formation algorithm for rational information agents. In: *Proceedings of the 2nd International Conference on MultiAgent Systems*, (pp. 157–164).
27. Korilis, Y., & Lazar, A. (1995). Architecting non-cooperative networks. *IEEE Journal on Selected Areas of Communication*, 13(7), 1241–1251.
28. Koutsoupias, E., & Papadimitriou, C. (1999). Worst-case equilibria. In: *In Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science*, (pp. 404–413).

29. Kraus, S., Shehory, O., & Tasse, G. (2003). Coalition formation with uncertain heterogeneous information. In: *Proceedings of AAMAS*, (pp. 1–8).
30. Moyaux, T., Chaib-draa, B., & D'Amours, S. (2006). Supply chain management and multiagent systems: An overview. In B. Chaib Draa abd & J. Muller (Eds.), *Multiagent based supply chain management* (pp. 1–27). Berlin: Springer.
31. von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behaviour*. Princeton, NJ: Princeton University Press.
32. Ordeshook, P. (1986). *Game theory and political theory: An introduction*. Cambridge: Cambridge University Press.
33. Osborne, M. J., & Rubinstein, A. (1994). *A course in game theory*. Cambridge: MIT Press.
34. Rahwan, T., Michalak, T. P., Wooldridge, M., & Jennings, N. R. (2015). Coalition structure generation: A survey. *Artificial Intelligence*, doi:[10.1016/j.artint.2015.08.004](https://doi.org/10.1016/j.artint.2015.08.004).
35. Ray, D. (2007). *A game-theoretic perspective on coalition formation*. Oxford: Oxford University Press.
36. Ray, D., & Vohra, R. (1999). A theory of endogenous coalition structures. *Games and Economic Behavior*, 26, 286–336.
37. Rota, G. (1964). The number of partitions of a set. *American Mathematical Monthly*, 71(5), 498–504.
38. Roth, A. (1977a). Utility functions for simple games. *Journal of Economic Theory*, 16(2), 481–489.
39. Roth, A. (1977b). The Shapley value as a von Neumann-Morgenstern utility. *Econometrica*, 45(3), 657–664.
40. Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, 50, 97–109.
41. Shehory, O., & Kraus, S. (1996). A kernel oriented model for coalition formation in general environments: Implementation and results. In: *Proceedings of AAAI*, (pp. 134–140).
42. Thrall, R., & Lucas, W. (1963). N-person games in partition function form. *Naval Research Logistics Quarterly*, 10(1), 281–298.
43. Wittman, D. (1989). Why democracies produce efficient results. *Journal of Political Economy*, 97(6), 1395–1424.
44. Wittman, D. (2004). Efficiency of democracy. In: *The Encyclopedia of Public Choice*, (pp. 510–513). Springer, Berlin.
45. Yi, S. (1997). Stable coalition structures with externalities. *Games and Economic Behavior*, 20, 201–237.
46. Yi, S. (2003). Endogenous formation of economic coalitions: A survey on the partition function approach. In: *The endogenous formation of economic coalitions*, (pp. 80–127). Edward Elgar, London, UK.