Coalition formation in serial dictatorships

Ben Plaut[1], Caleb Donovick[1], and Reyna Hulett[1]

[1]CS Department, Stanford University, Stanford, CA 94305.
Email:{bplaut, donovick, rmhulett}@stanford.edu
video link: https://www.youtube.com/watch?v=7JKvA6YK5jQ&feature=youtu.be

Abstract

In a serial dictatorship, agents are processed sequentially and assigned to their favorite remaining item on their turn. We introduce and analyze a class of coalition formation games in variants of the serial dictatorship mechanism. Rather than agents being processed in a random order, the order is determined in a principled way by the coalition structure. We demonstrate that these games are non-hedonic by showing that an agent's utility depends not only on the members of her coalition, but on the entire coalition structure. We primarily analyze the strategic behavior of selfish agents in these models. First, we consider a model where agents have no knowledge of other agents' preferences, and as such base decisions only on coalition sizes. We then consider a model in which agents have full knowledge of other agents' preferences, but only have utility for their top choice, and are only assigned if would they receive their top choice. For both of these models, we describe how an agent should choose which coalition to join, and when an agent should prevent others from joining her coalition. This leads to a characterization of equilibria and stable configurations. Finally, we explain why reasoning about full knowledge of preference lists is difficult. We wonder if aspects of such reasoning are in fact NP-hard, and leave this as an open question.

1 Introduction

Sometimes, success depends more on who your allies are than on who *you* are. Maybe you're a judge hoping for an appointment by the next president; or you work at Facebook and your cousin just happens to be Mark Zuckerberg; or you're looking for a team for a tug-of-war tournament. When you choose your allies, you don't know what team will end up winning, or who will become politically powerful or rich (maybe it will be you!), but the outcome most certainly will affect you. With such limited information, how should you choose to align yourself?

In this report, we analyze coalition formation in variants of the serial dictatorship mechanism. The primary distinction from the original mechanism is that, instead of agents being ordered randomly, a random dictator is chosen from among the agents, and this dictator and their coalition get assigned first. Equivalently, we can consider choosing a random coalition to go first, weighted by the coalition's size. In the examples above, the random dictator (coalition) would correspond to who gets elected president or becomes a rich CEO (or which tug-of-war team wins). Crucially, the identity of the dictator is not known during coalition formation, since otherwise they would have no incentive to join coalitions.

More formally, we study the following model for coalition formation in serial dictatorships (CFSD).

1.1 General model

Let A be a set of agents who are to be assigned to items in X, where $|A| = |X| = n$. Coalitions have no effect when $n \leq 2$, so we assume $n > 2$ throughout the paper. Each agent i has a ranking R_i over X. Define a *coalition* to be any non-empty subset of these agents. Let a *coalition structure* be a set of coalitions such that each agent is in exactly one coalition, i.e., a partitioning of the agents into coalitions.

Our general model, a variant of the serial dictatorship mechanism, consists of the following four steps:

1. *Ranking submission.* Each agent i submits a ranking R_i.
2. *Coalition formation.* A coalition structure $\mathcal{C}$ is determined by some mechanism.
3. *Global ordering.* A random dictator d is chosen from among the agents. The dictator's coalition $c(d)$ goes first, in a uniformly random ordering, followed by all the other agents in a uniformly random ordering.
4. *Serial assignment.* Agents are processed in the ordering determined by step 3, where each agent i is assigned her top remaining choice under R_i when it is her turn, or is unassigned if no ranked choice remains.

The global ordering and serial assignment steps are identical in all models considered in this report. We primarily consider three variables in this model: whether agents know each other's rankings, the number of items ranked, and whether an agent can prevent others from joining her coalition. Section 3 analyzes the model where agents have no knowledge of each other's preferences and all agents submit a full, strict ranking over X. Section 4 presents results for the model where agents know each other's preferences but each agent's ranking consists only of their top choice and they have no utility for any other assignment. For each of these models, we consider both the case where an agent can join whichever coalition she desires, and the case where the coalition's current members have the option to reject her request to join. Finally, Section 5 discusses the complications that arise when agents can have utility for more than one outcome *and* general knowledge of preferences.

2 Related Work

The highly combinatorial, probabilistic, and non-convex nature of CFSD makes it fairly unique. Combinatorial auctions are somewhat similar to CFSD as agents have preferences over a partitioning of resources. While it may appear fairly easy to model CFSD as a combinatorial auction with each agent bidding their expected utility for each coalition structure, even determining an agents expected utility for a given structure is itself, in the general case, a combinatorial problem. Further, it is unclear how solutions to combinatorial auctions correspond to stable structures in CFSD, which, combined with the apparent intractability of determining expected utility, makes work from combinatorial auction theory generally inapplicable.

Hedonic games are coalition formation games in which each player's utility is defined by the members of her coalition. We will demonstrate that CFSD is not such a game as each player's utility is a probability distribution defined not only by the members of their own coalition but by the global coalition structure. However, we explore many ideas that are similar to those in the hedonic literature.

2.1 Hedonic Games

Dreze and Greenberg [1] introduced the concept of hedonic coalitions in resource allocation games and explored the Pareto optimal coalition structures along with notions of stability, specifically individual stability [1] and contractual individual stability [2]. The purely hedonic model was formalized in 1998 by Bogomolnaia and Jackson [2]. In their work they explored four types of stability: individual stability, contractual individual stability, core stability [3], and Nash stability [4]. They also proved the necessary conditions for stability. Banerjee, Konishi and Sönmez [3] showed that the core of coalition formation games are generally empty without very strong properties. Ballester and Coralio showed that finding stable structures given arbitrary preferences is generally a NP-compete problem [4]. One might think that this would imply that finding stable structures in CFSD would also be NP-compete, however, the utility gained from possible coalition structures is highly structured (because of the entanglement with the serial dictatorship mechanism) which makes it difficult to reduce arbitrary hedonic games to CFSD.

[1] no player can switch into a coalition while which improve herself and not hurt the coalition she is switching into, a concept we used in what call exclusive formation games

[2] same as individually stability except that player must not hurt the coalition that she is leaving

[3] no set of players can opt out of the game to form a better coalition than the one they are currently in; this does not make much sense in our CFSD, as preferences are over the whole structure not just the agent's coalition

[4] no player can arbitrarily switch into a coalition and make herself better, something we considered in every model

3 Coalition formation with unknown preferences

In this section we consider coalition formation, under the assumption that agents have no knowledge of each other's preferences. Let $c(i)$ be the coalition to which agent i belongs, and let $s(i)$ be agent i's spot in the resulting ordering, where the best possible spot is 0 and the worst is $n-1$.

3.1 Expected utility

We begin with an exact calculation an agent's expected utility given any utility function over her preferences, and any arbitrary coalition structure. This will also demonstrate that coalition formation games are not hedonic, as the resulting expression can depend on the entire coalition structure.

Define $\mathbb{I}_{Rd} : X \to \{0, 1\}$ by

$$\mathbb{I}_{Rd}(x) = \begin{cases} 1 & \text{if } xRd \text{ where } R \text{ is some binary relation} \\ 0 & \text{otherwise} \end{cases}$$

Expected utility can be calculated as the sum over ℓ of the expected utility given being in spot ℓ times the probability of being in spot ℓ for $\ell \in \{0, 1, ..., n-1\}$. The expected utility given being in spot ℓ is the sum over k of the probability of getting choice k given spot ℓ times the expected utility of choice k. Note given spot ℓ a agent in the worst case will receive their ℓth choice. Now without loss of generality assume $X = \{0, 1, ..., n-1\}$ and i's preference list, R_i, is $(0, 1, ..., n-1)$ so i's kth choice is $k-1$. Let a_i be the assignment given to i (a random variable dependent on $s(i)$, R_i, and on the coalition structure $\mathcal{C}$). Finally, let $U_i(a_i)$ be the utility agent i gets from assignment a_i.

$$\begin{aligned} \mathbb{E}\big[U_i(a_i)\big] &= \sum_{\ell=0}^{n-1} \Big(\Pr(s(i) = \ell) \cdot \mathbb{E}\big[U_i(a_i \mid s(i) = \ell)\big] \Big) \\ &= \sum_{\ell=0}^{n-1} \Big(\Pr(s(i) = \ell) \cdot \sum_{k=0}^{n} \big(\Pr(a_i = k \mid s(i) = \ell) \cdot U_i(k) \big) \Big) \\ &= \sum_{\ell=0}^{n-1} \Big(\Pr(s(i) = \ell) \cdot \sum_{k=0}^{\ell} \big(\Pr(a_i = k \mid s(i) = \ell) \cdot U_i(k) \big) \Big) \end{aligned}$$

We next calculate $\Pr(s(i) = \ell)$. This is dependent on $\mathcal{C}$ and can be expressed as:

$$\begin{aligned} \Pr(s(i) = \ell) &= \mathbb{I}_{>\ell}(|c(i)|) \cdot \frac{|c(i)|}{n} \cdot \frac{1}{|c(i)|} + \sum_{c \in \mathcal{C} \setminus c(i)} \mathbb{I}_{\leq \ell}(|c|) \cdot \frac{|c|}{n} \cdot \frac{1}{n - |c|} \\ &= \mathbb{I}_{>\ell}(|c(i)|) \cdot \frac{1}{n} + \sum_{c \in \mathcal{C} \setminus c(i)} \mathbb{I}_{\leq \ell}(|c|) \cdot \frac{|c|}{n(n - |c|)} \end{aligned}$$

The indicator function is necessary, since $\Pr(s(i) = \ell) = 0$ if $d \in c(i)$ and $|c(i)| \leq \ell$. Similarly, $\Pr(s(i) = \ell) = 0$ if $d \notin c(i)$ and $|c(d)| > \ell$. For the sake of brevity we shall leave $\Pr(s(i) = \ell)$ for the remainder of the proof.

If we assume that all agents' preferences are uniformly distributed over the set of permutations of $(0, 1, ..., n-1)$, we can define i's probability of getting her kth choice given spot ℓ as the probability that a random permutation of elements from $\{0, 1, ..., n-1\}$ of length ℓ contains elements $\{0, 1, ..., k-1\}$ and not k[5]. However, since order doesn't matter, we will consider combinations instead.

By definition, there are $\binom{n}{\ell}$ ℓ length combinations drawn from $\{0, 1, ..., n-1\}$. The number of ℓ length combinations containing $\{0, 1, ..., k-1\}$ and not k can be counted as the number of $\ell - k$ length combinations

[5] The uniform distribution of preference induces uniform distribution of assignments. This is because when it is the first agent's turn to pick they will pick their first choice which is uniform distributed over X. Now by induction, given some uniformly selected $Y \subset X$ where Y are the choices of first $j-1$ agents, we see that agent j chooses the top choice from R_j which is not already in Y. As both are uniform, all $x \in X$ are equally likely to be in $X \setminus Y$ and to be top remaining choice in R_j. Hence agent j has a uniform probability of selecting any given $x \in X$. Therefore uniformly random preferences induce uniformly random assignments.

drawn from $\{0, 1, ..., n-1\} \setminus \{0, 1, ..., k\}$ which is simply $\binom{n-k-1}{\ell-k}$. Hence the probability of i getting her kth choice given spot ℓ is $\frac{\binom{n-k-1}{\ell-k}}{\binom{n}{\ell}}$. Now we have

$$\begin{aligned}\mathbb{E}\big[U_i(a_i)\big] &= \sum_{\ell=0}^{n-1}\Bigg(\Pr(s(i)=\ell)\cdot\sum_{k=0}^{\ell}\big(\Pr(a_i=k)\cdot U_i(k)\big)\Bigg)\\ &= \sum_{\ell=0}^{n-1}\Bigg(\Pr(s(i)=\ell)\cdot\sum_{k=0}^{\ell}\big(\frac{\binom{n-k-1}{\ell-k}}{\binom{n}{\ell}}\cdot U_i(k)\big)\Bigg)\end{aligned}$$

Thus given an agent's utility function and a coalition structure, we can in principle exactly compute her expected utility. However, our expression for $\mathbb{E}\big[U_i(a_i)\big]$ is quite unwieldy, and consequently not much help if we desire a clean analysis of strategies in coalition formation games.

3.2 Expected position in the ordering

Next, we calculate an exact expression for an agent i's expected position in the ordering under this model. Although the resulting expression is far from elegant, it is wieldy enough to be of use in an analysis of strategies in these games. In particular, we will use this expression in our proofs in Section 3.3 and Section 3.4.

Let $\alpha(i) = \mathbb{E}[|c(d)| \mid d \notin c(i)]$ for brevity. We handle $|c(i)| = n$ as a special case, since $\alpha(i)$ is undefined when $\Pr(d \notin c(i)) = 0$. Then for $|c(i)| < n$,

$$\begin{aligned}\mathbb{E}\big[s(i) \mid |c(i)| < n\big] &= \Pr(d \in c(i)) \cdot \mathbb{E}[s(i) \mid d \in c(i)] + p(d \notin c(i)) \cdot \mathbb{E}[s(i) \mid d \notin c(i)]\\ &= \frac{|c(i)|}{n}\Big(\frac{|c(i)|-1}{2}\Big) + \frac{n-|c(i)|}{n}\Big(\alpha(i) + \frac{n-\alpha(i)-1}{2}\Big)\\ &= \frac{|c(i)|^2 - |c(i)|}{2n} + \frac{(n-|c(i)|)(n+\alpha(i)-1)}{2n}\\ &= \frac{|c(i)|^2 - |c(i)| + n^2 + n\alpha(i) - n - |c(i)|n - |c(i)|\alpha(i) + |c(i)|}{2n}\\ &= \frac{|c(i)|^2 + n^2 + \alpha(i)(n-|c(i)|) - n(|c(i)|+1)}{2n}\end{aligned}$$

For $|c(i)| = n$, we have

$$\mathbb{E}\big[s(i) \mid |c(i)| = n\big] = \Pr(d \in c(i)) \cdot \mathbb{E}[s(i) \mid d \in c(i)] = 1 \cdot \frac{n-1}{2} = \frac{n-1}{2}$$

Observe that we can recover the $|c(i)| = n$ result from the more general expression, since the term containing the undefined $\alpha(i)$ drops out. Thus we will use the more general expression going forward.

The $\alpha(i)$ term gives a clear demonstration of the dependence of these games on the entire coalition structure. Intuitively, if the all of the agents not in $c(i)$ are in a single coalition c', then $|c'| = n - |c(i)|$ agents are guaranteed to be ahead of i if $d \notin c(i)$. On the other hand, if all agents not in $c(i)$ are in singleton coalitions, then only one spot will be taken by the dictator's coalition.

In the next section, we formally define the first coalition formation game we consider, and present our results for that game.

3.3 Free coalition formation

The Free Coalition Formation Game consists solely of each agent i choosing an integer from 1 to n inclusive, where each integer corresponds to a coalition. For each integer k chosen by at least one agent, the set of all agents who chose k forms an element of the resulting coalition structure, a coalition $c \subseteq A$. Agents cannot be prevented from joining a coalition in this version of the game.

Since we assume in Section 3 that agents have no knowledge of other agents' preferences, all they can do is attempt to minimize their spot in the ordering. Therefore for the remainder of Section 3, we assume each agent i receives utility $U_i(s(i)) = n - s(i)$. Note that this defines a fixed-sum game.

3.3.1 Strategies for free coalition formation

We now describe strategies for selfish agents in the Free Coalition Formation Game, ending with a characterization of the unique dominant strategy equilibrium.

Lemma 1. *Let c_1 and c_2 be two coalitions with $|c_2| \geq |c_1|$. Then any member of c_1 can strictly decrease her expected spot in the ordering by switching to c_2, assuming no other agents switch coalitions.*

Proof. Let i be a member of c_1. Let $|c_1| = k$ and $|c_2| = k + \delta - 1$ ($k + \delta$ if i switches) for some $\delta \in \mathbb{N}^+$. Let $\alpha_{c_1}(i)$ and $\alpha_{c_2}(i)$ be the values of $\alpha(i)$ for $i \in c_1$ and $i \in c_2$ respectively. We compare i's expected spot in the ordering if she remains in c_1 to her expected spot if she switches, and confirm that switching to c_2 yields a lower expected spot in the ordering.

We handle $k + \delta = n$ as a special case in order to prevent division by zero.

Case 1: $k + \delta = n$. Since only i switches, every other agent must already be in c_2. This implies that $k = 1$ and $\alpha_{c_1}(i) = n - 1$. Therefore

$$\begin{aligned}
\mathbb{E}\big[s(i) \mid i \in c_1\big] - \mathbb{E}\big[s(i) \mid i \in c_2\big] &= \frac{k^2 + n^2 + \alpha_{c_1}(i)(n-k) - n(k+1)}{2n} - \frac{n-1}{2} \\
&= \frac{1^2 + n^2 + (n-1)(n-1) - n(1+1) - n(n-1)}{2n} \\
&= \frac{n^2 - 3n + 2}{2n}
\end{aligned}$$

Thus since $n > 2$, we have $\mathbb{E}\big[s(i) \mid i \in c_1\big] - \mathbb{E}\big[s(i) \mid i \in c_2\big] > 0$. Therefore $\mathbb{E}\big[s(i) \mid i \in c_1\big] > \mathbb{E}\big[s(i) \mid i \in c_2\big]$ implying that i's expected spot is strictly decreased by switching to c_2.

Case 2: $k + \delta < n$. This is the case where the analysis is substantial. We have

$$\begin{aligned}
&\mathbb{E}\big[s(i) \mid i \in c_1\big] - \mathbb{E}\big[s(i) \mid i \in c_2\big] \\
&= \frac{k^2 + n^2 + \alpha_{c_1}(i)(n-k) - n(k+1)}{2n} - \frac{(k+\delta)^2 + n^2 + \alpha_{c_2}(i)(n-(k+\delta)) - n((k+\delta)+1)}{2n} \\
&= \frac{\big[k^2 - (k+\delta)^2\big] + \big[\alpha_{c_1}(i)(n-k) - \alpha_{c_2}(i)(n-k-\delta)\big] + \big[n(k+1+\delta) - n(k+1)\big]}{2n} \\
&= \frac{-2k\delta - \delta^2 + \big[\alpha_{c_1}(i)(n-k) - \alpha_{c_2}(i)(n-k-\delta)\big] + \delta n}{2n}
\end{aligned}$$

Recall that $\alpha(i)$ is the expected size of the dictator's coalition, given that the dictator is not in i's coalition. So

$$\alpha(i) = \sum_{c \neq c(i)} \Pr(d \in c \mid d \notin c(i)) \cdot |c| = \sum_{c \neq c(i)} \frac{|c|}{n - |c(i)|} \cdot |c| = \sum_{c \neq c(i)} \frac{|c|^2}{n - |c(i)|}$$

$$\begin{aligned}
\alpha_{c_1}(i)(n-k) - \alpha_{c_2}(i)(n-k-\delta) &= \Big[(n-k) \sum_{c \neq c_1} \frac{|c|^2}{n-k}\Big] - \Big[(n-k-\delta) \sum_{c \neq c_2} \frac{|c|^2}{n-(k+\delta)}\Big] \\
&= \sum_{c \neq c_1} |c|^2 - \sum_{c \neq c_2} |c|^2 \\
&= \Big[(k+\delta-1)^2 + \sum_{c \neq c_1, c_2} |c|^2\Big] - \Big[(k-1)^2 + \sum_{c \neq c_1, c_2} |c|^2\Big] \\
&= (k+\delta-1)^2 - (k-1)^2 \\
&= 2k\delta + \delta^2 - 2\delta
\end{aligned}$$

We now plug this into our expression for $\mathbb{E}\big[s(i) \mid i \in c_1\big] - \mathbb{E}\big[s(i) \mid i \in c_2\big]$.

$$\begin{aligned}
\mathbb{E}\big[s(i) \mid i \in c_1\big] - \mathbb{E}\big[s(i) \mid i \in c_2\big] &= \frac{-2k\delta - \delta^2 + 2k\delta + \delta^2 - 2\delta + \delta n}{2n} \\
&= \frac{\delta n - 2\delta}{2n}
\end{aligned}$$

Since $\delta > 0$ and $n > 2$, $\mathbb{E}\big[s(i) \mid i \in c_1\big] - \mathbb{E}\big[s(i) \mid i \in c_2\big] > 0$ in this case as well, which completes the proof. Note that plugging in $\delta = n-1$ here yields the expression that was derived in case 1. We also observe that the advantage of switching depends only on the total number of agents and the difference between the coalition sizes, not their absolute size. □

Corollary 3.0.1. *The change in utility from switching from c_1 to c_2 is zero if $|c_2| = |c_1| - 1$, negative if $|c_2| < |c_1| - 1$, and positive if $|c_2| \geq |c_1|$.*

Proof. Suppose $|c_2| = |c_1| - 1$. Then after the switch, $|c_1| = |c_2| - 1$. By symmetry, undoing the switch must yield the same change in utility as the original switch. Since the combined utility change from switching and undoing the switch must be zero, switching and undoing the switch must each yield zero change in utility.

Suppose $|c_2| < |c_1| - 1$. Then after switching to c_2, $|c_1| \geq |c_2|$. Then Lemma 1 implies that undoing the switch must increase the agent's utility; therefore the switch must have decreased the agent's utility.

Finally, if $|c_2| \geq |c_1|$, an application of Lemma 1 completes the proof. □

Theorem 3.1. *Belonging to the largest coalition is the unique dominant strategy for all agents in the Free Coalition Formation Game.*

Proof. This is an immediate consequence of Lemma 1. While any agent i is not in the strictly largest coalition, there must be another coalition of at least equal size, and she can always strictly increase her utility by switching to that coalition. She continues to do so until she belongs to the largest coalition. This is true regardless of the rest of the coalition structure. □

Corollary 3.1.1. *All agents choosing the same coalition is the unique dominant strategy equilibrium.*

3.4 Exclusive coalition formation

We would also like to analyze a model wherein coalitions can choose whether to allow new members to join. To do so, we must first define a new coalition formation game. In contrast to the Free Coalition Game, which was a one-shot game, this will be a sequential game.

We assume that each agent is initially in a coalition just with herself. In each round, we iterate through the agents in an arbitrary order that is the same for all rounds. Each agent, on her turn, may request to leave her current coalition and join a new one. She is accepted into the new coalition if all the existing members would benefit by adding the agent, i.e., if each member has a higher utility for the new coalition structure induced by accepting the request. We run rounds repeatedly until no agent proposes to join a new coalition.

An agent may only request to join coalition c given current coalition structure $\mathcal{C}$ if she has not previously requested to join c at a time when $\mathcal{C}$ was also the coalition structure. In other words, an agent can request to join the same coalition multiple times, even if rejected, but only if something has changed in the rest of coalition structure. This is because whether a request is accepted could in principle depend on the entire coalition structure.

Since there are a finite number of possible coalition structures and a finite number of coalitions, this game will always terminate. Also, note that the criteria for acceptance into a coalition (e.g., unanimity, majority, etc.) does not matter, because the agents of a coalition are indistinguishable if we disregard preferences, and thus should all benefit or not benefit together.

A strategy for an agent consists of a request (or pass) for every coalition structure. We continue to use the same utility function of $U_i(s(i)) = n - s(i)$. When we refer to an agent's utility at a point during the game, we mean her utility if the game were to terminate with that coalition structure.

3.4.1 Strategies for exclusive coalition formation

Lemma 1 gives conditions under which an agent wants to join a coalition. Lemma 1 still applies to exclusive coalition formation, but now we also need to know the conditions under which a request to join is accepted.

Lemma 2. *If an agent currently in c_1 requests to join c_2, c_2 will accept if and only if $|c_2| - |c_1| < \frac{n}{2} - 1$.*

Proof. The proof is similar that that of Lemma 1. Let i be an agent currently in c_2, and let j be the agent requesting to join c_2. Since we disregard preferences, all agents in c_2 will be affected in the same way. Therefore it suffices to show that i would benefit.

Let $|c_2| = k_2$ ($k_2 + 1$ if j is accepted), and let $|c_1| = k_1$ ($k_1 - 1$ if j is accepted). Let $\alpha_{c_1}(i)$ and $\alpha_{c_2}(i)$ be the values of $\alpha(i)$ for $j \in c_1$ and $j \in c_2$, respectively. We compare i's expected spot in the ordering if j is accepted to her expected spot if j is rejected. We again handle $k_2 + 1 = n$ as a special case to prevent division by zero.

Case 1: $k_2 + 1 = n$. In this case, initially $c_1 = \{j\}$ and $c_2 = A \backslash \{j\}$, so $\alpha_{c_1}(i) = 1$. Therefore

$$\begin{aligned}
\mathbb{E}\big[s(i) \mid j \in c_1\big] - \mathbb{E}\big[s(i) \mid j \in c_2\big] &= \frac{k_2^2 + n^2 + \alpha_{c_1}(i)(n - k_2) - n(k_2 + 1)}{2n} - \frac{n-1}{2} \\
&= \frac{(n-1)^2 + n^2 + 1 \cdot (n - (n-1)) - n((n-1)+1) - n(n-1)}{2n} \\
&= \frac{n^2 - 2n + 1 + n^2 + n - n + 1 - n^2 - n^2 + n}{2n} \\
&= \frac{2-n}{2n}
\end{aligned}$$

Since $n > 2$, $\mathbb{E}\big[s(i) \mid j \in c_1\big] - \mathbb{E}\big[s(i) \mid j \in c_2\big] < 0$. Therefore $\mathbb{E}\big[s(i) \mid j \in c_1\big] < \mathbb{E}\big[s(i) \mid j \in c_2\big]$, meaning that i's expected spot in the order is strictly increased by allowing j to join. Therefore j is rejected in this case.

Case 2: $k_2 + 1 < n$. In this case,

$$\begin{aligned}
&\mathbb{E}\big[s(i) \mid j \in c_1\big] - \mathbb{E}\big[s(i) \mid j \in c_2\big] \\
&= \frac{k_2^2 + n^2 + \alpha_{c_1}(i)(n - k_2) - n(k_2+1)}{2n} - \frac{(k_2+1)^2 + n^2 + \alpha_{c_2}(i)(n - (k_2+1)) - n((k_2+1)+1)}{2n} \\
&= \frac{\big[k_2^2 - (k_2+1)^2\big] + \big[\alpha_{c_1}(i)(n-k_2) - \alpha_{c_2}(i)(n - k_2 - 1)\big] + \big[n(k_2+2) - n(k_2+1)\big]}{2n} \\
&= \frac{-2k_2 - 1 + \big[\alpha_{c_1}(i)(n-k_2) - \alpha_{c_2}(i)(n-k_2-1)\big] + n}{2n}
\end{aligned}$$

Recall that $\alpha(i) = \sum_{c \neq c(i)} \frac{|c|^2}{n - |c(i)|}$. We proceed along the same lines as in the proof of Lemma 1. The key difference that is there for this lemma, i remains in the same coalition, so the limits of both sums are $c \neq c_2$.

$$\begin{aligned}
\alpha_{c_1}(i)(n-k_2) - \alpha_{c_2}(i)(n-k_2-1) &= \Big[(n-k_2) \sum_{c \neq c_2 | j \in c_1} \frac{|c|^2}{n-k_2}\Big] - \Big[(n-k_2-1) \sum_{c \neq c_2 | j \in c_2} \frac{|c|^2}{n-(k_2+1)}\Big] \\
&= \sum_{c \neq c_2 | j \in c_1} |c|^2 - \sum_{c \neq c_2 | j \in c_2} |c|^2 \\
&= \Big[k_1^2 + \sum_{c \neq c_1, c_2} |c|^2\Big] - \Big[(k_1-1)^2 + \sum_{c \neq c_1, c_2} |c|^2\Big] \\
&= 2k_1 - 1
\end{aligned}$$

Therefore:

$$\mathbb{E}\big[s(i) \mid j \in c_1\big] - \mathbb{E}\big[s(i) \mid j \in c_2\big] = \frac{-2k_2 - 1 + 2k_1 - 1 + n}{2n} = \frac{n + 2(k_1 - k_2 - 1)}{2n}$$

This implies that $\mathbb{E}\big[s(i) \mid j \in c_1\big] > \mathbb{E}\big[s(i) \mid j \in c_2\big]$ if and only if $n + 2(k_1 - k_2 - 1) > 0$. Therefore accepting j strictly decreases i's expected position if and only if $k_2 - k_1 < n/2 - 1$, as required. □

We say that a coalition structure is individually stable if there is no agent who can increase her utility without any other agents switching coalitions.

Theorem 3.2. *Any individually stable coalition structure of the Exclusive Coalition Formation Game consists of either (1) exactly two coalitions, one having size at least* $\frac{3n-2}{4}$*, or (2) a single coalition containing all agents.*

Proof. First, we show that no coalition structure with more than two coalitions is individually stable. Suppose there are at least three coalitions, and let c_1, c_2, and c_3 be three coalitions. Without loss of generality assume $1 \leq |c_1| \leq |c_2| \leq |c_3|$. Note that $|c_3| \leq n-2$, since $|c_1| + |c_2| + |c_3| \leq n$.

Then: $(|c_3| - |c_2|) + (|c_2| - |c_1|) \leq (n - 2 - |c_2|) + (|c_2| - 1) = n - 3$. Therefore at least one of $(|c_3| - |c_2|)$ and $(|c_2| - |c_1|)$ is at most $(n-3)/2$.

Suppose $|c_2| - |c_1| \leq (n-3)/2 < (n-2)/2 = n/2 - 1$. Then c_2 will accept an agent from c_1 which requests to join, by Lemma 2. By Lemma 1, any agent in c_1 would prefer to join c_2. Since there is an agent who wishes to join a coalition which would accept her, this coalition structure is not individually stable. The case for $|c_3| - |c_2| \leq (n-3)/2$ follows similarly. This shows that no coalition structure with more than two coalitions is individually stable.

Now suppose there are exactly two coalitions, c_1 and c_2. Without loss of generality assume $|c_2| \geq |c_1|$. Then any agent from c_1 always would prefer to switch to c_2, again by Lemma 1. Suppose $|c_2| < \frac{3n-2}{4}$. Then $|c_1| = n - |c_2| > \frac{n+2}{4}$.

Therefore we have $|c_2| - |c_1| < \frac{3n-2}{4} - \frac{n+2}{4} = \frac{2n-4}{2} = n/2 - 1$. Thus if $|c_2| < \frac{3n-2}{4}$, c_2 is willing to accept an agent from c_1, so any individually stable coalition structure must have $|c_2| \geq \frac{3n-2}{4}$.

If $|c_2| \geq \frac{3n-2}{4}$, we have $|c_2| - |c_1| \geq \frac{3n-2}{4} - \frac{n+2}{4} = n/2 - 1$, so by Lemma 2, c_2 will not accept any agents from c_1. By Corollary 3.0.1, no agent from c_2 would wish to switch to c_1, since $|c_1| < |c_2|$. Therefore a coalition structure with exactly two coalitions is individually stable if and only if one has size at least $3n/4$.

Finally, a single coalition is indeed individually stable, because Corollary 3.0.1 implies that any deviation strictly decreases an agent's utility. □

We define the utility of a move by agent i to be the utility of agent i after the move minus the utility of agent i before the move. We call an agent myopic if she always makes the move with highest utility. We call an agent weakly myopic if she makes a positive utility move whenever one exists, and never makes a negative utility move.

Lemma 3. *If all agents are weakly myopic, the coalition structure upon termination must be individually stable.*

Proof. We know that the game always terminates: let T be the time at which it terminates. Let $\mathcal{C}_t$ be the coalition structure at time t.

Suppose $\mathcal{C}_T$ is not individually stable. Then there exists an agent i and coalition c such that i could improve her utility by switching to c, and c would accept her. If i has never before requested to join c under $\mathcal{C}_T$, then switching to c is a legal positive utility move. Therefore i has at least one positive utility move, so by assumption, she must make a positive utility move. Therefore the game does not terminate at time T, a contradiction.

Thus i must have previously requested to join c under the same coalition structure $\mathcal{C}_T$. Assume that i first requested to join c under $\mathcal{C}_T$ at time t'. Therefore $\mathcal{C}_T = \mathcal{C}_{t'}$. Since c is willing to accept i and had not accepted i under $\mathcal{C}_T$ before by assumption, i successfully joined c at time t'.

Define a function $\phi(t)$ by $\phi(t) = \sum_{c' \in \mathcal{C}_t} |c'|^2$. It is trivial that an agent choosing to pass on her turn leaves $\phi(t)$ unchanged. Suppose at time t, an agent moves from c_1 to c_2, where $k_1 = |c_1|$ and $k_2 = |c_2|$ at time t (so $|c_1| = k_1 - 1$ and $|c_2| = k_2 + 1$ at time $t+1$). Then

$$\begin{aligned}\phi(t+1) - \phi(t) &= \sum_{c' \in \mathcal{C}_{t+1}} |c'|^2 - \sum_{c' \in \mathcal{C}_t} |c'|^2 \\ &= \Big[(k_1 - 1)^2 + (k_2 + 1)^2 + \sum_{c' \in \mathcal{C}_{t+1} \setminus \{c_1, c_2\}} |c'|^2\Big] - \Big[k_1^2 + k_2^2 + \sum_{c' \in \mathcal{C}_t \setminus \{c_1, c_2\}} |c'|^2\Big]\end{aligned}$$

Since all coalitions other than c_1 and c_2 are unchanged by this move, we have

$$\sum_{c' \in \mathcal{C}_{t+1} \setminus \{c_1, c_2\}} |c'|^2 = \sum_{c' \in \mathcal{C}_t \setminus \{c_1, c_2\}} |c'|^2$$

and therefore

$$\begin{aligned}
\phi(t+1) - \phi(t) &= (k_1 - 1)^2 + (k_2 + 1)^2 - k_1^2 - k_2^2 \\
&= k_1^2 - 2k_1 + 1 + k_2^2 + 2k_2 + 1 - k_1^2 - k_2^2 \\
&= 2(k_2 - k_1 + 1)
\end{aligned}$$

Thus $\phi(t+1) - \phi(t) > 0$ whenever $k_2 \geq k_1$, and $\phi(t+1) - \phi(t) < 0$ only if $k_2 < k_1 - 1$. Then by Corollary 3.0.1, $\phi(t+1) - \phi(t) > 0$ for any positive utility move, and $\phi(t+1) - \phi(t) < 0$ only for negative utility moves.

Since $\mathcal{C}_T = \mathcal{C}_{t'}$, $\phi(T) = \phi(t')$. By assumption, agent i switching to c at time t' is a positive utility move, so $\phi(t'+1) - \phi(t') > 0$. Thus there must exist some t'' where $t' < t'' < T$ where $\phi(t''+1) - \phi(t'') < 0$. But this implies that an agent made a negative utility move, which is a contradiction.

Therefore, if each agent makes a positive utility whenever possible, and never makes a negative utility move, the coalition structure upon termination must be individually stable. □

Theorem 3.3. *If all agents are weakly myopic, then the Exclusive Coalition Formation Game always terminates with exactly two coalitions, one with size exactly* $\lfloor 3n/4 \rfloor$.

Proof. We now show by induction that no coalition ever reaches size greater than $\lfloor 3n/4 \rfloor$.

Initially all coalitions have size 1, so the base case is trivially satisfied. For the inductive step, suppose that at time t, no coalition has size greater than $\lfloor 3n/4 \rfloor$. Let c_1 be the largest coalition at time $t+1$. Then $|c_1| \leq \lfloor 3n/4 \rfloor$.

Suppose $|c_1| \leq \lfloor 3n/4 \rfloor - 1$. Since at most one agent changes coalitions per round, $|c_1| \leq \lfloor 3n/4 \rfloor$ at time $t+1$, and we are done.

Therefore assume $|c_1| = \lfloor 3n/4 \rfloor$. Since $\lfloor 3n/4 \rfloor \geq \frac{3n-2}{4}$, any other coalition c_2 has size at most $n - |c_1| \leq \frac{n+2}{4}$. Thus, $|c_1| - |c_2| \geq \frac{3n-2}{4} - \frac{n+2}{4} = n/2 - 1$ for any $c_2 \neq c_1$. Therefore $|c_1|$ would not accept an agent from any coalition, implying that $|c_1| = \lfloor 3n/4 \rfloor$ at time $t+1$. This completes the induction and shows that no coalition ever reaches size greater than $|c_1| = \lfloor 3n/4 \rfloor$.

By Lemma 3, the coalition structure upon termination must be individually stable. Then by Theorem 3.2, we must terminate with a coalition of size at least $\frac{3n-2}{4}$. Since $\lfloor 3n/4 \rfloor$ is the smallest integer that is at least $\frac{3n-2}{4}$, we must terminate with a coalition of size at least $\lfloor 3n/4 \rfloor$. Since by the above induction, no coalition ever reaches size greater than $\lfloor 3n/4 \rfloor$, we must terminate with a coalition of size exactly $\lfloor 3n/4 \rfloor$.

Therefore there must be at least two coalitions upon termination. By Theorem 3.2, any individually stable coalition structure has at most two coalitions. Thus we must terminate with exactly two coalitions, which completes the proof. □

It is worth noting that there are cases where making a negative utility move can be better in the long term, although they require the rest of the agents to have unintuitive strategies. For example, suppose there are two coalitions c_1 and c_2 where $|c_1| = n/4$ and $|c_2| = 3n/4$. Suppose agent i is in c_1. A request from i to join c_2 would be rejected, so i can either remain in c_1, or leave to form a coalition consisting of just herself. Suppose all of the agents in c_2 have the following strategy: join i's new coalition if i leaves c_1, otherwise do nothing. Thus if i remains in c_1, she will end up with a coalition of size $n/4$, but if she leaves, she will end up with a coalition of much larger size, which is better for her in the end.

4 Petulant child model

We now move on to a second class of coalition formation games. Unlike in Section 3, the petulant child model assumes that agents have complete knowledge of other agents' preferences. However, we make the simplifying assumption that all agents care only about their odds of getting their top choice, know every other agent's top choice, *and* submit rankings containing only this top choice. Thus if an agent's top choice is not available on her turn, she is unassigned.

Therefore agents only want to maximize the probability of being the first of the agents with the same top choice. This is similar to a small child who does not get their favorite toy from a toybox. Instead of agreeing to play with their second favorite toy, they may choose to throw a tantrum and receive zero utility. While it may seem unlikely that such an agent would be capable of strategic thought, we shall analyze their optimal behavior anyway.

We will also assume that the top choice submitted by agent i in R_i and used in the assignment step is the same top choice known to all other agents. Under this restriction, the mechanism is truthful, since agents have no utility for any choice other than their top choice.

In the following, let M_i be the set of agents with the same top choice as agent i. We will assume $M_i > 1$, since otherwise i will get her top choice no matter what, and has no incentive to participate in forming coalitions.

4.1 Expected utility

Since agents now have knowledge of each other's preferences, we must study the expected utility directly and not use position as a proxy. Let $\beta(i) = \Pr(M_i \cap c(d) = \emptyset \mid d \notin c(i))$; essentially, this is the probability that i has a chance of getting her top choice given that her coalition isn't chosen. Then the probability i gets her top choice is:

$$\begin{aligned}
&\Pr(d \in c(i))\frac{1}{|M_i \cap c(i)|} + \Pr(d \notin c(i))\beta(i)\frac{1}{|M_i|} \\
=&\frac{|c(i)|}{n|M_i \cap c(i)|} + \frac{(n-|c(i)|)\beta(i)}{n|M_i|} \\
=&\frac{|c(i)||M_i| + \beta(i)|M_i \cap c(i)|(n-|c(i)|)}{n|M_i||M_i \cap c(i)|}
\end{aligned}$$

4.2 Free coalition formation

First, we will study optimal behavior and equilibria using the Free Coalition Formation Game introduced in Section 3.3.

4.2.1 Strategies for free coalition formation

We now describe optimal strategies for selfish agents in the Free Coalition Formation Game and petulant child model, ending with a characterization of the Nash equilibria.

Lemma 4. *For an agent i choosing among coalitions which all have competitors (other members of M_i) or all do not, i maximizes her expected utility by seeking the highest ratio of coalition size to number of competitors after i joins $\frac{|c|+1}{|M_i \cap c|+1}$, assuming no other agents switch coalitions.*

Proof. Consider an agent i deciding whether to switch from coalition c_1 to c_2. i's change in utility from switching will be

$$\begin{aligned}
&\frac{(|c_2|+1)|M_i| + \beta_{c_2}(i)(|M_i \cap c_2|+1)(n-|c_2|-1)}{n|M_i|(|M_i \cap c_2|+1)} - \frac{|c_1||M_i| + \beta_{c_1}(i)|M_i \cap c_1|(n-|c_1|)}{n|M_i||M_i \cap c_1|} \\
=&\frac{(|c_2|+1)|M_i||M_i \cap c_1| - |c_1||M_i|(|M_i \cap c_2|+1)}{n|M_i||M_i \cap c_1|(|M_i \cap c_2|+1)} + \frac{\beta_{c_2}(i)(n-|c_2|-1) - \beta_{c_1}(i)(n-|c_1|)}{n|M_i|} \\
=&\frac{(|c_2|+1)|M_i \cap c_1| - |c_1|(|M_i \cap c_2|+1)}{n|M_i \cap c_1|(|M_i \cap c_2|+1)} + \frac{\left(\sum_{c \neq c_2 \cup \{i\}} |c| \text{ if } M_i \cap c = \emptyset\right) - \left(\sum_{c \neq c_1 \setminus i} |c| \text{ if } M_i \cap c = \emptyset\right)}{n|M_i|} \\
=&\frac{(|c_2|+1)|M_i \cap c_1| - |c_1|(|M_i \cap c_2|+1)}{n|M_i \cap c_1|(|M_i \cap c_2|+1)} + \frac{(|c_1|-1 \text{ if } M_i \cap c_1 \setminus i = \emptyset) - (|c_2| \text{ if } M_i \cap c_2 = \emptyset)}{n|M_i|}
\end{aligned}$$

We want to consider two cases: either i is switching between coalitions where neither has competitors ($M_i \cap c_1 \backslash i = M_i \cap c_2 = \emptyset$) or where both have competitors ($M_i \cap c_1 \backslash i, M_i \cap c_2 \neq \emptyset$). In the first case, $M_i \cap c_2 = M_i \cap c_1 \backslash i = \emptyset$, i.e., i is considering two coalitions where no one shares her top choice, so i should switch if

$$\frac{|c_2| + 1 - |c_1|}{n} + \frac{|c_1| - 1 - |c_2|}{n|M_i|} \geq 0$$
$$|c_2| + 1 \geq |c_1|$$

As we would expect, i should always switch to a larger coalition if neither coalition contains a competitor for her top choice. However, if $M_i \cap c_1 \backslash i, M_i \cap c_2 \neq \emptyset$, i.e., both coalitions have competitors, i should switch if

$$(|c_2| + 1)|M_i \cap c_1| - |c_1|(|M_i \cap c_2| + 1) \geq 0$$
$$\frac{|c_2| + 1}{|M_i \cap c_2| + 1} \geq \frac{|c_1|}{|M_i \cap c_1|}$$

So interestingly, if both have competitors, what matters is the ratio of size to number of competitors. This matches what we saw if both coalitions lack competitors, since the denominators in the above expression will just be 1.

Thus for any pair of coalitions (both with or both without competitors), i maximizes her utility by maximizing the ratio of coalition size to number of competitors. By transitivity, i therefore maximizes her utility in the same way when choosing among any number of such coalitions. □

This clean result unfortunately does not hold, as we will see below, if only one of the two coalitions has competitors. Intuitively, if one coalition lacks competitors, then i still has a chance of getting her top choice when that coalition is picked, regardless of whether she is in that coalition, and the odds of this happening depend on the size of that coalition.

Lemma 5. *For an agent i choosing among coalitions only some of which have competitors, i maximizes her expected utility by seeking the highest ratio of coalition size to number of competitors after i joins $\frac{|c|+1}{|M_i \cap c|+1}$ but with an incentive to avoid coalitions without competitors equal to $\frac{|c|}{|M_i|}$, assuming no other agents switch coalitions.*

Proof. Using the same expression for the change in i's utility as in Lemma 4, if $M_i \cap c_1 \backslash i \neq \emptyset$ and $M_i \cap c_2 = \emptyset$, i.e., i is considering switching from a coalition with competitors to one without, i should only switch if

$$\frac{(|c_2| + 1)|M_i \cap c_1| - |c_1|}{n|M_i \cap c_1|} - \frac{|c_2|}{n|M_i|} \geq 0$$
$$|c_2||M_i||M_i \cap c_1| + |M_i||M_i \cap c_1| - |c_1||M_i| - |c_2||M_i \cap c_1| \geq 0$$
$$|c_2| \geq \frac{|c_1||M_i| - |M_i||M_i \cap c_1|}{|M_i||M_i \cap c_1| - |M_i \cap c_1|}$$

or more intuitively, recalling $|M_i \cap c_2| = 0$,

$$|c_2| + 1 - \frac{|c_1|}{|M_i \cap c_1|} - \frac{|c_2|}{|M_i|} \geq 0$$
$$\frac{|c_2| + 1}{|M_i \cap c_2| + 1} \geq \frac{|c_1|}{|M_i \cap c_1|} + \frac{|c_2|}{|M_i|}$$

So in essence we still care about the ratio of coalition size to number of competitors, but c_1 gets a benefit. Specifically, if i stays with c_1, i still has a $1/|M_i|$ chance of winning if c_2 is chosen as well as a $1/|M_i \cap c_1|$ chance if c_1 is chosen, whereas if i switches she has a $1/|M_1 \cap c_2| + 1 = 1$ chance of winning if c_2 is chosen but no chance if c_1 is. On the other hand, if $M_i \cap c_1 \backslash i = \emptyset$ and $M_i \cap c_2 \neq \emptyset$, i.e., only c_2 has competitors, we should just get the reverse of the above equation. But as a sanity check, i should switch if

$$\frac{|c_2| + 1 - |c_1|(|M_i \cap c_2| + 1)}{n(|M_i \cap c_2| + 1)} + \frac{|c_1| - 1}{n|M_i|} \geq 0$$
$$|c_2||M_i| + |M_i| - |c_1||M_i||M_i \cap c_2| - |c_1||M_i| + (|M_i \cap c_2| + 1)(|c_1| - 1) \geq 0$$

or more intuitively, recalling $|M_i \cap c_1| = 1$,

$$\frac{|c_2| + 1}{|M_i \cap c_2| + 1} - |c_1| + \frac{|c_1| - 1}{|M_i|} \geq 0$$
$$\frac{|c_2| + 1}{|M_i \cap c_2| + 1} + \frac{|c_1| - 1}{|M_i|} \geq \frac{|c_1|}{|M_i \cap c_1|}$$

which does indeed match the previous inequality. Again, we care about the ratio of coalition size to number of competitors, but this time c_2 gets a benefit. Specifically, if i switches to c_2, i still has a $1/|M_i|$ chance of winning if c_1 is chose as well as a $1/|M_i \cap c_2| + 1$ if c_2 is, whereas if i stays she has a $1/|M_1 \cap c_1| = 1$ chance of winning if c_1 is chosen but no chance if c_2 is.

Thus for any pair of coalitions where exactly one has competitors, i maximizes her utility by maximizing the ratio of coalition size to number of competitors but with a benefit of coalition size over $|M_i|$ to avoiding the coalition without competitors. By transitivity, i therefore maximizes her utility in the same way when choosing among any number of such coalitions. □

Note that for the specific case where c_1 is the coalition with just i, it's always better for i to switch to c_2; if c_2 has no competitors then i should switch into the larger coalition, and if it does have competitors then the above inequality reduces to $\frac{|c_2|+1}{|M_i \cap c_2|+1} \geq 1$ which is always true.

Now we can reason about Nash equilibria in the free coalition formation game; a coalition structure is a Nash equilibrium (i.e., produced by a strategy profile which is a Nash equilibrium) if no agent would strictly benefit by switching to another coalition. The conditions under which agents would benefit by switching are detailed above. For simplicity, we will assume agents i such that $|M_i| = 1$ don't join any coalitions, since they get their top choice regardless of the coalition structure. We will therefore ignore these agents in the subsequent characterization of equilibria.

Clearly there can be many possible Nash equilibria. The coalition of everyone is an equilibrium, since as we noted above, it's always better for an agent to be in a coalition with others than to be on her own. However, it's also an equilibrium to have any number of coalitions with identical top-choice profiles, where every top choice is represented multiple times per coalition, since if any agent switched they would end up in a coalition with a worse size to competitors ratio.

Thus, we cannot give a clean description of all Nash equilibria. However, we will give one straightforward necessary condition ("top choice profile containment"), and then a less straightforward set of sufficient conditions, for a coalition structure to be an equilibrium.

Theorem 4.1. *If a coalition structure with coalitions $|c_1| \leq |c_2| \leq \cdots \leq |c_k|$ is a Nash equilibrium, it must obey* top choice profile containment. *That is, for all $\ell \in \{1, \ldots, k-1\}$, the top-choice profile of c_ℓ must be contained in the top-choice profile of $c_{\ell+1}$. Equivalently, for any agent $i \in c_\ell$, $|M_i \cap c_\ell| \leq |M_i \cap c_{\ell+1}|$.*

Proof. Assume to the contrary that for some coalition c_ℓ with agent i, we had $|M_i \cap c_\ell| > |M_i \cap c_{\ell+1}|$. Since $|c_\ell| \leq |c_{\ell+1}|$, if i switched she would be in a strictly larger coalition with no more competitors than before. Per the conditions above, if c_ℓ and $c_{\ell+1}$ both or neither have competitors, this would strictly improve i's utility by increasing the size to competitors ratio. Otherwise, c_ℓ must have competitors and $c_{\ell+1}$ must not, so i would strictly improve her utility by switching if $\frac{|c_{\ell+1}|+1}{|M_i \cap c_{\ell+1}|+1} = |c_{\ell+1}| + 1$ is strictly larger than

$\frac{|c_\ell|}{|M_i \cap c_\ell|} + \frac{|c_{\ell+1}|}{|M_i|}$. Since c_ℓ has competitors and $|c_\ell| \leq |c_{\ell+1}|$,

$$\begin{aligned}\frac{|c_\ell|}{|M_i \cap c_\ell|} + \frac{|c_{\ell+1}|}{|M_i|} &\leq \frac{|c_\ell|}{2} + \frac{|c_{\ell+1}|}{2} \\ &< \frac{|c_{\ell+1}|+1}{2} + \frac{|c_{\ell+1}|}{2} = |c_{\ell+1}| + 1\end{aligned}$$

Thus in any circumstances, switching would strictly improve i's utility, which contradicts that the initial coalition structure was a Nash equilibrium. Thus no coalition structure can be an equilibrium unless it obeys top choice profile containment. □

The containment requirement illuminates some other interesting conditions for Nash equilibria.

Corollary 4.1.1. *If two coalitions in an equilibrium are the same size, they must be identical in terms of top choice profile.*

Corollary 4.1.2. *If any M_i is contained entirely in one coalition, it must be the strictly largest coalition.*

Corollary 4.1.3. *k, the number of coalitions, is at most equal to the largest number of agents with the same top choice.*

Proof. c_1 must contain at least one agent with some top choice t, but by containment, so must all the coalitions in the chain. Thus the number of coalitions is at most the number of agents with top choice t, which is at most equal to the largest number of agents with the same top choice.

Note that this limit on k is achievable, for instance, if all agents have the same top choice. In that case, all coalition structures are Nash equilibria, because agents always have utility $1/n$ since they only get their top choice if they are chosen as the dictator. Thus each agent being in her own coalition is an equilibrium, and $k = n =$ the number of agents sharing the same top choice. □

However, note that the containment requirement is *not* sufficient to guarantee that the structure is an equilibrium. For instance, if some $M_i \subsetneq c_k$, i could strictly benefit from switching to c_{k-1} if $|c_{k-1}| + 1 > \frac{|c_k| + |c_{k-1}|}{|M_i|}$. As a specific example, this inequality holds for the coalition structure where c_1 has top-choice profile $\{1\}$ and c_2 has $\{1, 2, 2, 2\}$, so this structure is not an equilibrium even though it obeys the containment rule.

We can give an inelegant characterization of the necessary and sufficient conditions for an equilibrium. We will start with the containment rule,

$$\forall \ell \in \{1, \ldots, k-1\} \; \forall i \in c_\ell \; |M_i \cap c_\ell| \leq |M_i \cap c_{\ell+1}| \tag{1}$$

and figure out what other conditions must be met to guarantee that no agent would benefit from switching.

First, consider an agent i who is the only agent of M_i in $c(i)$. The containment rule guarantees i will not want to switch to a coalition with no members of M_i, since it implies all such coalitions are strictly smaller and thus switching would not increase the size of the coalition i belongs to. However, it must also be the case that i would not benefit from switching to a coalition containing at least one member of M_i. Thus, if $c_{\ell-1} \cap M_i = \emptyset$ and $c_\ell \cap M_i = \{i\}$,

$$\forall m > \ell \quad \frac{|c_m| + 1}{|M_i \cap c_m| + 1} \leq |c_\ell| - \frac{|c_\ell| - 1}{|M_i|} \tag{2}$$

This is necessary because otherwise i would benefit from switching to some larger coalition. It is also sufficient to guarantee that no other agent $j \in M_i$ with no competitors in their coalition $c(j)$ would benefit from switching to a coalition with competitors. This is because $|c(j)| \geq |c_\ell|$ by definition. Thus j wouldn't strictly benefit by switching to c_ℓ because $\frac{|c_\ell|+1}{|M_i \cap c_\ell|+1} = \frac{|c_\ell|+1}{2} \leq \frac{|c(j)|+1}{2} \leq |c(j)| - \frac{|c(j)|-1}{|M_i|}$. Nor would j strictly benefit by switching to c_m for $m > \ell$ because $\frac{|c_m|+1}{|M_i \cap c_m|+1} \leq |c_\ell| - \frac{|c_\ell|-1}{|M_i|} \leq |c(j)| - \frac{|c(j)|-1}{|M_i|}$.

Finally, we need to ensure that no agent i with competitors in $c(i)$ would benefit from switching. If ℓ is the highest index such that $c_\ell \cap M_i = \emptyset$ and ℓ' is the highest index such that $c'_\ell \cap M_i \leq 1$,

$$\forall m > \ell' \quad |c_\ell| + 1 \leq \frac{|c_m|}{|M_i \cap c_m|} + \frac{|c_\ell|}{|M_i|} \tag{3}$$

This must hold because otherwise some $i \in c_m$ would benefit from switching to c_ℓ. However, it also ensures no agent i with competitors in $c(i)$ would strictly benefit from switching to a coalition c without competitors, since by definition $|c| \leq |c_\ell|$ and thus the desired inequality holds for any pair of coalitions, one with no members of M_i and one with multiple.

It must also be the case that no i with competitors in $c(i)$ would benefit from switching to a coalition with competitors. Thus for all $c_\ell \neq c_m$ with $|c_\ell \cap M_i| \geq 2$ and $|c_m \cap M_i| \geq 1$,

$$\frac{|c_m| + 1}{|M_i \cap c_m| + 1} \leq \frac{|c_\ell|}{|M_i \cap c_\ell|} \tag{4}$$

Thus, Equations 1 - 4 constitute necessary and sufficient conditions for a coalition structure to be a Nash equilibrium.

4.3 Exclusive coalition formation

We will now study optimal behavior using the Exclusive Coalition Formation Game introduced in Section 3.4. However, in this setting, it actually matters what method we use for voting on new coalition members. Since agents now have knowledge of preferences, we no longer expect the vote to be unanimous.

4.3.1 Strategies for exclusive coalition formation

We have already calculated above the conditions under which an agent would want to switch coalitions. However, for the exclusive coalition formation game, we also want to know when an agent would accept a new member of the coalition.

Lemma 6. *Agent i always benefits by adding $j \notin M_i$ to $c(i)$ (although not strictly if $M_i \subset c(i)$).*

Proof. Consider the difference in agent i's utility if j joins $c(i)$ versus if he remains in $c(j)$:

$$\begin{aligned}
&\frac{(|c(i)|+1)|M_i| + \beta_{c(i)}(i)|M_i \cap c(i)|(n - |c(i)| - 1)}{n|M_i||M_i \cap c(i)|} - \frac{|c(i)||M_i| + \beta_{c(j)}(i)|M_i \cap c(i)|(n - |c(i)|)}{n|M_i||M_i \cap c(i)|} \\
&= \frac{1}{n|M_i \cap c(i)|} + \frac{\left(\sum_{c \neq c(i) \cup \{j\}} |c| \text{ if } M_i \cap c = \emptyset\right) - \left(\sum_{c \neq c(i)} |c| \text{ if } M_i \cap c = \emptyset\right)}{n|M_i|} \\
&= \frac{1}{n|M_i \cap c(i)|} + \frac{(|c(j)| - 1 \text{ if } M_i \cap c(j) \backslash j = \emptyset) - (|c(j)| \text{ if } M_i \cap c(j) = \emptyset)}{n|M_i|} \\
&= \frac{1}{n|M_i \cap c(i)|} + \frac{-1 \text{ if } M_i \cap c(j) = \emptyset}{n|M_i|} \geq 0
\end{aligned}$$

Thus it is always advantageous to add an agent with a different top choice to your coalition. (Note that the above expression equals zero if and only if $|M_i \cap c(i)| = |M_i|$.) This makes sense because agents only care about being the first among those with the same top preference (and if all of M_i is in the same coalition, they always have a $1/|M_i|$ chance of winning). ☐

Lemma 7. *Agent i benefits by adding $j \in M_i$ to $c(i)$ if and only if $M_i \cap c(j) = \{j\}$ and $\frac{|c(i)|+1}{|M_i \cap c(i)|+1} + \frac{|c(j)|-1}{|M_i|} \geq \frac{|c(i)|}{|M_i \cap c(i)|}$ (or if $c(i) \subset M_i$).*

Proof. Consider the difference in agent i's utility if $j \in M_i$ remains in $c(j)$ versus if he joins $c(i)$:

$$\frac{|c(i)||M_i| + \beta_{c(j)}(i)|M_i \cap c(i)|(n - |c(i)|)}{n|M_i||M_i \cap c(i)|} - \frac{(|c(i)|+1)|M_i| + \beta_{c(i)}(i)(|M_i \cap c(i)|+1)(n-|c(i)|-1)}{n|M_i|(|M_i \cap c(i)|+1)}$$
$$= \frac{(|M_i \cap c(i)|+1)|c(i)| - |M_i \cap c(i)|(|c(i)+1|)}{n|M_i \cap c(i)|(|M_i \cap c(i)|+1)} + \frac{\beta_{c(j)}(i)(n-|c(i)|) - \beta_{c(i)}(n-|c(i)|-1)}{n|M_i|}$$
$$= \frac{|c(i)| - |M_i \cap c(i)|}{n|M_i \cap c(i)|(|M_i \cap c(i)|+1)} + \frac{\left(\sum_{c \neq c(i)} |c| \text{ if } M_i \cap c = \emptyset\right) - \left(\sum_{c \neq c(i) \cup \{j\}} |c| \text{ if } M_i \cap c = \emptyset\right)}{n|M_i|}$$
$$= \frac{|c(i)| - |M_i \cap c(i)|}{n|M_i \cap c(i)|(|M_i \cap c(i)|+1)} + \frac{(|c(j)| \text{ if } M_i \cap c(j) = \emptyset) - (|c(j)| - 1 \text{ if } M_i \cap c(j) \backslash j = \emptyset)}{n|M_i|}$$
$$= \frac{|c(i)| - |M_i \cap c(i)|}{n|M_i \cap c(i)|(|M_i \cap c(i)|+1)} - \frac{|c(j)| - 1 \text{ if } M_i \cap c(j) \backslash j = \emptyset}{n|M_i|}$$

Clearly if $M_i \cap c(j) \backslash j \neq \emptyset$, then this expression is non-negative, and i suffers by adding j to $c(i)$ (though if $|c(i)| = |M_i \cap c(i)|$, this expression is less than or equal to zero in any case). That is to say, it's always bad to add someone sharing your top choice, except possibly if adding them would remove the last competitor from another coalition. In that case,

$$= \frac{|c(i)| - |M_i \cap c(i)|}{n|M_i \cap c(i)|(|M_i \cap c(i)|+1)} - \frac{|c(j)| - 1}{n|M_i|}$$
$$= \frac{|c(i)||M_i| - |M_i||M_i \cap c(i)| - |c(j)||M_i \cap c(i)|(|M_i \cap c(i)|+1) + |M_i \cap c(i)|(|M_i \cap c(i)|+1)}{n|M_i||M_i \cap c(i)|(|M_i \cap c(i)|+1)}$$

so i benefits by adding j if j is the only member of $c(j)$ sharing i's top choice and

$$0 \geq |c(i)||M_i| - |M_i||M_i \cap c(i)| - |c(j)||M_i \cap c(i)|(|M_i \cap c(i)|+1) + |M_i \cap c(i)|(|M_i \cap c(i)|+1)$$
$$|c(j)| \geq 1 + \frac{|M_i|\left(|c(i)| - |M_i \cap c(i)|\right)}{|M_i \cap c(i)|(|M_i \cap c(i)|+1)}$$

or perhaps more intuitively,

$$0 \geq \frac{|c(i)| + |c(i)||M_i \cap c(i)|}{|M_i \cap c(i)|(|M_i \cap c(i)|+1)} - \frac{|M_i \cap c(i)| + |c(i)||M_i \cap c(i)|}{|M_i \cap c(i)|(|M_i \cap c(i)|+1)} - \frac{|c(j)| - 1}{|M_i|}$$
$$\frac{|c(i)|+1}{|M_i \cap c(i)|+1} + \frac{|c(j)|-1}{|M_i|} \geq \frac{|c(i)|}{|M_i \cap c(i)|}$$

That is, i would want $j \in M_i$ to join $c(i)$ if and only if j is the only member of $c(j)$ sharing i's top choice, and the advantages to i of the $i/|M_i|$ chance of winning if $c(j) \backslash j$ is chosen outweighs the disadvantage of adding j. ☐

Theorem 4.2. *Under the voting rule where an agent j is accepted into coalition c unless the utiltiy of* every *member of c would decrease, the individually stable outcomes of the Exclusive Coalition Formation Game are identical to the equilibria of the Free Coalition Formation Game.*

Proof. Under this voting rule, no request is ever denied. Since agents always benefit by adding a new member not sharing their top choice, the only way a request could be denied is if all members of the voting coalition $c(i)$ share the same top choice as the requester j. However, we saw above that if $c(i) \subset M_i$ then all agents are neutral to adding j, so they would still accept the request to join.

Recall that a coalition structure is an equilibrium of the Free Coalition Formation Game if and only if no agent would benefit by switching to another coalition. Also, a coalition structure is individually stable in the Exclusive Coalition Formation Game if and only if no agent would benefit by switching to another coalition *and* that coalition would accept them. However, since coalitions always accept requests under this rule, the two conditions are identical, and equilibria of the Free Coalition Formation Game are identical to individually stable structures in the Exclusive Coalition Formation Game. ☐

Of course, we would expect to see fairly different results for other voting rules (e.g., majority rule or unanimity to accept), since not all agents in a coalition would vote the same way.

5 Coalition formation with knowledge of preferences

Reasoning about full preferences with complete information is incredibly difficult. Even calculating expected utility of agent i given coalition structure $\mathcal{C}$ at least naively requires calculating the results for all possible orderings. This is because even a single swap in the ordering can have large cascading effects on the resulting assignments. For example consider the following preference profiles: $\{R_1 = (x_1, x_2, x_3, x_4), R_2 = (x_2, x_1, x_3, x_4), R_3 = (x_2, x_3, x_4, x_1,), R_4 = (x_3, x_4, x_1, x_2)\}$, if the agents go in order (3,2,4,1) we end up with the following mapping of resources to agents $\{x_2 : 3;\ x_1 : 2;\ x_3 : 4;\ x_4 : 1\}$, note that agent 1 received her last choice. Now if we make a single swap and go in order (2,3,4,1) then the resulting map is $\{x_2 : 2;\ x_3 : 3;\ x_4 : 4;\ x_1 : 1\}$, which resulted in agent 1 receiving her first choice.

Further, even if we could calculate the expected utility for an agent given a structure, we are left with a difficult non-convex optimization problem to determine the optimal coalition structure for a given player. Suppose we have a large number of players and we are trying to optimize the results for a given player i with $R_i = (x_1, x_2, ...x_n)$. Suppose i only has utility for her first choice. This example is not equivalent to the petulant child model, as other agents may have utility beyond their first choice, and regardless of utility, they are guaranteed to be assigned. Finally, suppose that all other agents rank x_1 third in their list.

When i's coalition has size at most two, all other agents look the same: they all increase the chance her coalition will be chosen first without increasing the chance that a member of her coalition will take her first choice. However, as the size of the coalition grows, i must reason about which agents are least likely to end up taking her first choice. This, in turn, depends on how likely it is that those agents first or second choices are available on their turn. The result is a complex web of probabilistic dependencies, potentially involving cascades of the sort described in the beginning of this section.

Clearly this reasoning is complicated even when i only cares about her top choice, and with more general utility functions we would only expect it to become more complex.

6 Conclusion

We have examined a number of models of CFSD and shown that these games are not hedonic, as each agent's utility is dependent on the global coalition structure and not just the membership of her coalition. We explored the necessary conditions for stability. We proved that, given a few assumptions about rational play, stable coalition structures will emerge in the zero-knowledge variants of the game. We also identified key elements of the petulant child model, such as the ratio of coalition size to number of competitors, and top choice profile containment.

6.1 Open Questions

In the version of the game where agents do not know each other's preferences, we assumed that the utility of agent i was $U_i(s(i)) = n - s(i)$. How does optimal behavior change if U_i is in some other form? Suppose all preferences are drawn from some distribution: what happens when agents do not know each other's preferences, but do know the distribution from which they are drawn? For agent i, is the size of her optimal coalition based on the probability of other agents having similar preferences? Another interesting avenue could be to vary how the dictator is selected: what happens when the dictator is drawn from a known distribution that is not uniform? In that model, agents who are more likely to be dictator have more value as coalition members. How does this affect the strategies of individual agents?

Given the hardness results in other combinatorial games [4, 5], we wonder if CFSD with full knowledge is similarly difficult. However, reducing any such games to CFSD is difficult because it is not clear how one can express arbitrary utility functions in coalition structures. This is because an agent's utility does not depend directly on the coalition structure; it depends on the results of the serial dictatorship. The coalition structure simply has side effects on the probability distribution that defines an agent's utility, which as we have explained, are hard to analyze.

Although we proved the existence of Nash equilibria in the petulant child model, the following questions remain: Will all games arrive at a Nash equilibrium if agents are allowed to switch their strategies indefinitely? Given a voting rule, is it possible for cycles between coalition structures to arise in the exclusive version of that model? If so, is there any voting rule which guarantees eventual convergence of strategies? We assumed agents are bound to the preferences they report during coalition formation; if this were not the case, when do agents have incentive to misrepresent their preferences?

References

[1] J. H. Dreze and J. Greenberg, "Hedonic coalitions: Optimality and stability," *Econometrica: Journal of the Econometric Society*, pp. 987–1003, 1980.

[2] A. Bogomolnaia and M. O. Jackson, "The stability of hedonic coalition structures," *Games and Economic Behavior*, vol. 38, no. 2, pp. 201–230, 2002.

[3] S. Banerjee, H. Konishi, and T. Sönmez, "Core in a simple coalition formation game," *Social Choice and Welfare*, vol. 18, no. 1, pp. 135–153, 2001.

[4] C. Ballester, "Np-completeness in hedonic games," *Games and Economic Behavior*, vol. 49, no. 1, pp. 1–30, 2004.

[5] V. Conitzer, T. Sandholm, and P. Santi, "Combinatorial auctions with k-wise dependent valuations," in *AAAI*, vol. 5, 2005, pp. 248–254.