# Self-enforcing Coalitions with Power Accumulation

Karl Jandoc and Ruben Juarez*

Department of Economics and UHERO, University of Hawaii

2424 Maile Way, Honolulu, HI 96822 (Corresponding Author: rubenj@hawaii.edu)

January 2016

**Abstract**

Agents endowed with power compete for a divisible resource by forming coalitions with other agents. The coalition with the greatest power wins the resource and divides it among its members. The agents' power increases according to their share of the resource.

We study two models of coalition formation where winning agents accumulate power and losing agents may participate in further coalition formation processes. An axiomatic approach is provided by focusing on variations of two main axioms: *self-enforcement*, which requires that no further deviation happens after a coalition has formed, and *rationality*, which requires that agents pick the coalition that gives them their highest payoff. For these alternative models, we determine the existence of stable coalitions that are *self-enforcing* and *rational* for two traditional sharing rules.

The models presented in this paper illustrate how power accumulation, the sharing rule, and whether losing agents participate in future coalition formation processes, shape the way coalitions will be stable throughout time.

**Keywords:** Coalition Formation, Power Accumulation, Self-enforcement.

**JEL Classification** C70 · D71

*Financial support from the AFOSR Young Investigator Program is greatly appreciated. Any errors or omissions are the authors' own. We are grateful for helpful comments by two outstanding referees, the associate editor, Daron Acemoglu, Anna Bogomolnaia, Vikram Manjunath, Hervé Moulin and seminar participants at the Econometric Society Conference in Mexico, Conference on Economic Design in Lund, Public Economic Theory Conference in Lisbon and the International Conference on Game Theory in Stony Brook.

# 1 Introduction

Recent developments in the coalition formation literature have focused on different notions of stability.[1] This is largely motivated by popular results in political economy which show that stability in the ruling coalition leads to more favorable welfare outcomes in the long run. These results show that stable coalitions have a higher "effective discount factor" that enables them to look at longer horizons for the consequences of their actions. For example, McGuire and Olson [24] and Olson [28] compare public policy outcomes of a long-lived to a short-lived ruling coalition and found better outcomes in the former. In the context of a principal-agent political economy model, Barro [5], Ferejohn [13], Persson, Roland and Tabellini [30, 31] also found that politicians who are likely to remain in office for an extended period of time also supported better policies. In contrast, Acemoglu, Golosov and Tsyvinski [3] disputes the finding that stable ruling groups lead to fewer economic distortions. This result happens when ruling coalitions can be rewarded or punished even after they have left power. Furthermore, there have been recent attempts to test the robustness of different stability notions by adding dimensions relating to power, resource sharing, and other institutional constraints. "Power"–which is measured in the literature in terms of wealth, military might or political influence–reflects the ability of a coalition to impose its will on the rest of the society. Another enrichment of recent coalition formation models involves how resources (i.e., prize or spoils) are divided within the winning coalition. Finally, variations in the institutional environment of the coalition formation process (in our case whether agents are killed or are able to survive in future rounds) will impinge on the possible coalitions that form.

In this paper, we present a dynamic model of coalition formation where farsighted agents are endowed with power and compete for a divisible resource by forming coalitions with other agents. The assumption of farsightedness assures that agents care about their payoff from being included in the final (or *limit*) coalition that forms rather than their immediate payoffs (Konishi and Ray [21], Chwe [10], Xue [40], Diamantoudi and Xue [11], Ray and Vohra [35].) If a formed coalition has sufficient power to win, its members will divide the resource among themselves and their power will increase according to the member's share of the prize.[2] We focus on two sharing rules of the

[1] For a recent survey of the growing literature on coalition formation, refer to Ray [33], Ray and Vohra [34], Bloch and Dutta [7].

[2] The canonical example of power accumulation comes from non-democratic societies, where typically a ruling coalition can perpetuate itself in authority because it can use the state's resources to consolidate and accumulate power and wealth. Although data is sparse, there is some evidence that some world leaders amassed personal wealth either due to connections or investments made possible by (sometimes corrupt

prize: sharing the prize in proportion to an agent's power in the winning coalition (*proportional sharing*) and sharing the prize equally among the winning agents (*equal sharing*). An institutional constraint emerging from the dynamic coalition formation model is the span of life of the agents. We focus on two polar cases: (1) when agents who are not part of the winning coalition do not participate in the future (*agents are killed*); and, (2) when all the agents remain in the game regardless of whether or not they are part of the winning coalition (*agents survive*).[3]

We employ an axiomatic approach to find rules that choose coalitions (which we call *transition correspondences*) that satisfy two main axioms (*self-enforcement* and *rationality*) for the four cases above (combinations of proportional vs equal sharing and agents are killed vs agents survive). In order to describe self-enforcement, in a dynamic model, it is possible that over time coalitions may disintegrate and new factions may form to overthrow the existing ruling coalition. Also, since agents are farsighted, they want to form coalitions that are stable from the moment of inception. Self-enforcement requires that no subcoalition of the winning coalition is powerful enough to encourage further deviations. It is a robustness property that ensures that the coalition that forms in round one never disintegrates afterwards. Rationality requires that agents pick the coalition that gives them their highest limit payoff among self-enforcing coalitions. Rationality is related to the traditional axiom of coalitional stability, where no coalition will have the incentive to deviate.[4]

In the next subsection, we provide an illustrative example of how our coalition formation game works and how sharing of the prize and whether agents are killed or are able to survive shape the coalitions that form.

---

practices) using state resources. For instance, there is some speculation that Teodoro Obiang Nguema Mbasogo of Equitorial Guinea has amassed a fortune of over $600 million in his oil-rich country since ascending to power in 1979 (see http://www.huffingtonpost.com/2013/11/29/richest-world-leaders_n_4178514.html). Power accumulation is natural in many coalition formation settings ranging from labor unions to military alliances between countries, where the players generate benefits by being together and become more powerful as time passes.

[3]There are valid reasons to believe that the decision to kill agents in the society may be endogenous, especially in certain political contests or military wars. For instance, Bueno de Mesquita, et al [9] asserts that "[w]hen the private benefits of office or coalition membership are large, people are more prepared to engage in horrendous acts of cruelty against others to ensure that their personal privileges are not lost". The size of these private benefits can be reflected by resources to be won in a coalition formation game. On the other hand, most of the biggest purges in the $20^{th}$ century are identified with specific personalities, e.g. Saddam Hussein, Adolf Hitler and Josef Stalin (Goode [14, 15]). To the extent that these purges are sometimes personality-driven, the decision to kill opponents is exogenous. While the case where the decision to kill agents is endogenous is certainly an appealing study, it is beyond the scope of this paper and we leave it for future research. In this paper, we assume that the environment where agents are killed or are able to survive is exogenous.

[4]This is related to immunity to group manipulations in models discussed by Bogomolnaia and Jackson [8], Ehlers [12], Juarez [20], Papai [29].

## 1.1 An illustrative example

Society is composed of 4 agents $\{1, 2, 3, 4\}$ with power profile $\pi = [1, 1.5, 8, 10]$. The game is played for an infinite number of periods and for each round the prize to be divided by the winning coalition is $I = 3$. An agent's share of the prize will be added to his power in the subsequent rounds of the game. A coalition's power is simply the sum of the agents' powers inside it. A winning coalition should have more than 50% of the total power in the society at any given round. Since agents are farsighted, they only care about their payoffs from being in the final (or limit) coalition.

Case 1: Non-winning agents are killed, proportional sharing.

First, no coalition of size 1 is winning. At the initial round (call this round 0) all winning coalitions of size 2 involves agent 4, i.e., $\{1, 4\}$, $\{2, 4\}$, $\{3, 4\}$. Suppose $\{3, 4\}$ forms in round 1. Power is added to agents 3 and 4 according to their share in the prize, that is, with proportional sharing agent 3 gets $\frac{8}{18}$ of the prize while agent 4 gets $\frac{10}{18}$. In round 1 after $\{3, 4\}$ forms, agents 1 and 2 are killed; agent 3's power is $8 + \frac{24}{18}$ while agent 4's power is $10 + \frac{30}{18}$. Left alone to themselves, agent 4 can now kill agent 3 since he has the higher power. Thus in round 2 and afterwards agent 4 will be the limit (singleton) coalition. Therefore, $\{3, 4\}$ is not self-enforcing because agent 3 will never agree to form this coalition since he will not be part of the limit coalition. This is true of all coalitions of size 2.

Also, no coalition of size 3 is self-enforcing. Since all size 3 coalitions involve either agent 3 or 4, at round 1 when we add power to the winning agents either agent 3 or 4 can kill the rest of the agents in the three person coalition. For instance, if $\{2, 3, 4\}$ forms and after sharing the prize, in round 1 agents 2's power is $1.5 + \frac{4.5}{19.5}$, agent 3's power is $8 + \frac{24}{19.5}$ and agent 4's power is $10 + \frac{30}{19.5}$. Therefore agent 4 can kill both agents 2 and 3 since he has a higher power than the others combined. Following this argument, a three person coalition will never form.

Therefore, the only self-enforcing coalition is the grand coalition $\{1, 2, 3, 4\}$ since no subcoalition has the incentive to deviate following the logical process outlined above.

Case 2: Non-winning agents are killed, equal sharing.

With the prize shared equally among agents in the winning coalition, the coalition $\{2, 3, 4\}$ will be self-enforcing. To see this, note that at round 1 after power has been added agent 2's power is now 2.5, agent 3's power is 9 and agent 4's power is 11. Thus, in contrast to the previous case, agent 4 will not have sufficient power to kill agents 2 and 3. Furthermore, $\{2, 3, 4\}$ cannot deviate to a coalition of size 2 since the higher powered agent in the 2-person coalition can kill the lower-powered agent. Hence, coalition $\{2, 3, 4\}$ can form in round 1 and will be stable forever.

Case 3: Non-winning agents survive, equal sharing.

When non-winning agents survive and the prize is shared equally, coalition $\{3, 4\}$ can be self-enforcing. Unlike in the previous two cases where the higher-powered agent can kill the lower-powered agent in a 2-person coalition, agent 4 in this case will not have sufficient power to deviate from $\{3, 4\}$. To see this, note that agents 1 and 2 will still survive even if $\{3, 4\}$ forms, and thus in round 1 (after $\{3, 4\}$ forms and power has accumulated) agent 1's power (1) plus agent 2's power (1.5) plus agent 3's power $(8 + 1.5)$ is now greater than agent 4's power $10 + 1.5$. Thus, agent 4 cannot deviate to a singleton winning coalition. Coalition $\{3, 4\}$ will form in round 1 and will be stable forever.

Case 4: Non-winning agents survive, proportional sharing.

A curious feature in this case is that coalitions will have the incentive to "jump" from one coalition to another when agents survive. Take the example when $\{1, 2, 3\}$ forms. At round 1 after power has accumulated with proportional sharing, agent 1's power is $1 + \frac{3}{10.5}$, agent 2's power is $1.5 + \frac{4.5}{10.5}$, agent 3's power is $8 + \frac{24}{10.5}$ while agent 4's power remains at 10. At round 2, agents 1 and 2 can ditch agent 3 and align themselves with agent 4 since they can get a higher share of the prize by doing so (note that agent 3's power is larger than agent 4's power at round 1). At round 3, faced with the same incentive, agents 1 and 2 will again move to coalition $\{1, 2, 3\}$ since these agents can get a relatively higher share by deviating to this coalition. This phenomenon is true for any winning coalition other than the grand coalition. This example illustrates the difficulty of finding self-enforcing coalitions in this case. In a later section, we introduce a class of priority mechanisms that satisfy the axioms.

## 1.2 Overview of the results

Section 3 of the paper studies the case where agents are killed. Proposition 1 describes the unique transition correspondence that satisfies self-enforcement, rationality and *scale invariance*[5] under proportional sharing. On the other hand, under equal sharing the class of self-enforcing, rational and scale invariant transition correspondences may

[5] A transition correspondence satisfies scale invariance if any scale in the power vector would not change the coalition chosen by the transition correspondence.

not exist. Proposition 2($i$) provides the largest class of coalition formation games where a self-enforcing and rational transition correspondence exists for equal sharing. In this domain of games, Proposition 2($ii$) shows the unique transition correspondence that meets self-enforcement, rationality and scale invariance. Roughly speaking, the transition correspondence characterized in Proposition 2($ii$) picks the smallest winning coalition where the powers of the agents are relatively equal, and is of size $2^k - 1$ for some natural number $k$.

Section 4 of the paper examines the case where agents survive. Proposition 3 characterizes the unique transition correspondence under equal sharing that satisfies self-enforcement, rationality, scale invariance and *independence of zeros*.[6] This transition correspondence picks the smallest winning coalitions of size $2^k$ for some natural number $k$. Proposition 3 also shows that self-enforcement, rationality, scale invariance and independence of zeros are not compatible under proportional sharing. The class of self-enforcing, rational and scale invariant transition correspondences under proportional sharing is large (see Proposition 4).

The paper is organized as follows. Section 2 describes the model. Sections 3 and 4 study self-enforcing coalitions when agents are killed and survive, respectively. Section 5 concludes. All proofs are in the Appendix.

## 1.3 Related literature

There are several ways by which "power" is treated in contemporary coalition formation literature. Piccione and Razin [32] examine how power relations determine the ranking of agents in society. The identity of the coalitions (as characterized by the power of agents within that coalition) determines the social order and thus the structure of society. Jordan [17] characterizes the core and the stable set in a class of coalitional games called "pillage games". In this model, wealth is allocated among the finite agents in the game. A reallocation of wealth among the agents is only made possible by using force. A power function, which is monotonically increasing in membership and the members' wealth, regulates the ability of the agents to use force. The coalition is then able to appropriate the wealth of other less powerful coalitions.

Although Jordan's model describes how power leads to wealth, it could also be possible that wealth creates power. In the traditional Olsonian (Olson [27]) coalition formation context, we can imagine a model where the wealthy, who are fewer in number relative to the poor, will be able to appropriate more political power (that is, sway

[6] A transition correspondence satisfies independence of zeros if agents with zero power do not affect the coalition chosen by the transition correspondence.

public policy in their favor) by being able to organize themselves with less cost. In this sense, wealth leads to power. In our model, we avoid this complication by implicitly assuming that wealth and power are interchangeable.

While newer coalition formation models are dynamic in nature, most of them are limited to a static distribution of power among the agents. We fill this gap by exploring the possibility that agents are able to accumulate power over time as long as they are part of a winning coalition.[7]

The manner in which coalitions divide the prize among member agents has also received attention not only in the coalition formation literature but also in related political economy studies. Acemoglu, et al [1] [hereafter AES] examine a coalition formation process where agents split the prize in proportion to an agent's relative power in the coalition. Bartling and von Siemens [6] characterize the emergence of equal sharing as an optimal solution to an incentive problem in a partnership. In this paper, we focus on describing the stable coalitions that emerge under these two popular and commonly used sharing rules–proportional sharing and equal sharing. While we assume that the sharing rule in this paper is fixed and exogenously given, an interesting and natural extension not covered here would be to endogenize the sharing rule. There are several strategies to achieve this, for instance, in the context of rent-seeking tournaments, Nitzan [25], Lee [22, 23] and Baik [4] contend that the long-run group sharing rule can fall anywhere from being undefined to extremely equal depending on the context of the mechanism to determine these rules (Nitzan [26]).

Finally, other factors have been built in newer models to capture the institutional environment of the coalition formation process. AES, for instance, considers an extreme non-democratic institutional setting wherein non-winning agents are killed. Sekeris [36], also in a non-democratic setting, does not allow the possibility that the ruling coalition eliminate opponents. In this paper, we investigate these two polar cases.

Our work is closely related to AES wherein farsighted agents are endowed with power and form coalitions with the goal of becoming the *ultimate ruling coalition.* In their paper, the winning coalition will split a given resource in proportion to its power and agents outside the winning coalition are killed. However, we extend AES on several fronts: First, we allow the possibility that non-winning agents survive throughout the whole coalition formation game.[8] Second, we allow the winning coalition to accumulate

[7]This was first articulated by Tullock [39], where he argues that since formal institutions are weak or absent regarding distribution and sharing of power, succession of leaders, and generating consensus, a ruling coalition ("junta") will degenerate into a dictatorship, that is, there will be power accumulation by one of the junta members until he becomes the sole ruler.

[8]In another paper, Acemoglu, et al [2], show that the results in the AES case extend even to situations where agents are not killed. The stable ruling coalition in this case is the "minimally winning coalition"

power over time. Third, while AES uses proportional sharing in their model, we also examine equal sharing where agents equally divide the prize among themselves.

The variations in the model that takes into account whether non-winning agents can participate in future coalition formation processes, and different ways of sharing the prize will provide a rich characterization of stable coalitions. Here are some of the main lessons:

- Although our coalition formation environment is fairly simple, some of our model's scenarios suggest that it would be difficult to find rules to pick coalitions that would be both rational and self-enforcing unless we restrict some of the game's parameters. As we will show later, these restrictions may be done on the powers of the agents or the possible coalitions that will form.
- The "rules of the game" of our model, described by the different variations in the sharing rules (i.e., equal vs. proportional) and the non-winning agents' life span (i.e. whether or not they are killed) have a large effect on the the existence of a rational and self-enforcing transition correspondence and therefore the shape of the coalitions that can form. For instance, under proportional sharing when non-winning agents are killed, there always exists a unique transition correspondence that meets self-enforcement and rationality. This will not be true, however, under equal sharing.
- There are particular coalitions sizes which are not likely to form under the different scenarios in our model. For instance, a coalition of size two can never form when agents are killed. This is because the agent with the highest power can always eliminate the one with the lowest power. In contrast, in other scenarios only certain sizes are allowed: those of size $2^k - 1, k \in \mathbb{N}$ under equal sharing and agents are killed and size $2^k, k \in \mathbb{N}$ under equal sharing and agents survive.

## 2 The model

Consider the set $N = \{1, \ldots, n\}$ of initial agents who are endowed with powers (e.g., wealth, guns) $\pi = [\pi_1, \ldots, \pi_n]$, respectively. A coalition $S$ is a subset of $N$, that is, $S \subseteq N$. The set of coalitions are all possible subsets of $N$, denoted by $2^N$. A coalition formation game is a pair $(S, \pi)$ where $S \subseteq N$ and $\pi \in \mathbb{R}_+^S$. The set of

which is the coalition with the smallest power among all winning coalitions. In their dynamic game, this corresponds to the case when all agents are eligible to vote on the proposed coalition (i.e., the democratic case). In the case where the voting power is limited to the ruling coalition but agents are not killed, the mapping in AES is modified to include all alternative coalitions not just the subset of the ruling coalition.

coalition formation games is denoted by **G**. We assume that power is additive, that is, the power of coalition $S$ is the sum of all powers of the agents inside the coalition, $\pi(S) = \sum_{i \in S} \pi_i$.[9] We denote by $\pi_S$ the restriction of the vector $\pi \in \mathbb{R}^N_+$ over coalition $S$. Let $1_S \in \mathbb{R}^S$ be the vector where all coordinates are equal to 1.

**Definition 1** *Given a game $(T, \pi)$, the* ***set of winning coalitions***[10] *is:*

$$W_{(T,\pi)} = \{S \subset T | \pi(S) > \pi(T \setminus S)\}$$

We fix a sharing rule that allocates the prize to the agents at every coalition formation game. Throughout the paper we devote special attention to two simple and commonly used sharing rules, equal sharing and proportional sharing (see Juarez [19]). That is, for the game $(T, \pi)$, let $\xi(T, \pi) \in \mathbb{R}^T$ be the vector of shares under the sharing rule $\xi$. Under **equal sharing (ES)**, $\xi = ES$, the share of agents $i \in T$ equals $\xi_i(T, \pi) = \frac{1}{|T|}$. On the other hand, under **proportional sharing (PR)**, $\xi = PR$, the share of agents $i \in T$ equals $\xi_i(T, \pi) = \frac{\pi_i}{\pi(T)}$.

Given the sharing rule $\xi$, let $\xi^N(T, \pi) \in \mathbb{R}^N$ be the embedded vector $\xi(T, \pi)$ into $\mathbb{R}^N$. That is, $\xi^N_i(T, \pi) = \xi_i(T, \pi)$ if $i \in T$ and $\xi^N_i(T, \pi) = 0$ if $i \in N \setminus T$.

The major task of this paper is to examine the stable coalitions that emerge under the two sharing rules discussed above. As we will see below, the type of stable coalitions will greatly vary depending on the sharing rule and whether agents survive or are killed.

## 2.1 Dynamic Coalition Formation

The game is played in discrete rounds. Let $t$, where $t = 0, 1, \dots,$ denote the time of the game. We define a *transition correspondence* that maps from the set of coalition formation games to a particular set of winning coalitions.

**Definition 2** *A* **transition correspondence** *is a continuous*[11] *correspondence $\phi$ : $\mathbf{G} \to 2^N$ such that $\forall$ $(X, \pi) \in \mathbf{G}$: $\phi(X, \pi) \subset W_{(X,\pi)}$.*

The transition correspondence $\phi$ selects all coalitions that could be winning at a given game $(S, \pi)$. The evolution of the coalition formation games at every round

[9] Juarez [19] considers a more general version where power can be any arbitrary monotonic function.

[10] This definition requires winning coalitions to have relative power larger than 50%. Our results below can be easily adapted to require winning coalitions to have relative power larger than $\alpha$, where $\alpha \geq 50\%$. This is discussed in Section 5.1.

[11] A correspondence is continuous if for any sequence of power vectors $\pi^1, \pi^2, \dots \to \pi^*$ where $S \in \phi(N, \pi^i)$ $\forall i$ and $S$ is winning in $\pi^*$, then $S \in \phi(N, \pi^*)$.

depends on the transition correspondence along with the conditions on whether agents survive or are killed in future rounds. The main task of this paper is to axiomatize $\phi$ in the four distinct scenarios that we describe below.

We denoted by $(S^t, \pi^t)$ the game at round $t$. We start with the game $(S^0, \pi^0)$ and look at the evolution of the game throughout time. We denote this evolution by $(S^0, \pi^0), (S^1, \pi^1), (S^2, \pi^2), \ldots$.

In this paper, we focus at the evolution of the game in two polar cases. First, we consider the case where agents are killed if they are not part of the winning coalition, this is described by the case where $S^t \in \phi(S^{t-1}, \pi^{t-1})$ for all time $t$. That is, only a group of agents who were winners at time $t-1$ could participate at time $t$. Second, we consider the case where agents survive regardless of whether they are part of the coalition in the previous round. This is described by $S^t = N$ for all time $t$. That is, the agents participating at every time do not change.[12]

We now describe the manner in which power evolves throughout time. First, we fix the prize $I > 0$ that is divided at every round. The manner by which the power of the agents accumulates will be affected by both the sharing rule $\xi$ and the transition correspondence $\phi$. That is, we assume that the prize $I$ is transformed to power in the same proportion. Formally, the power accumulation function for agent $i$ at time $t > 0$ equals:

$$\pi_i^t = \begin{cases} \pi_i^{t-1} + \xi_i(S^{t-1}, \pi^{t-1}) I & \text{if } i \in S^t \\ \pi_i^{t-1} & \text{if } i \in N \setminus S^t \text{ and agents survive} \end{cases}$$

Note that the power of the winning agent $i$ increases by $\xi_i(S^{t-1}, \pi^{t-1})$ regardless of whether agents are killed or survive. On the other hand, if non-winning agents survive, then their power does not increase. If agents are killed, they are removed from the game.

Implicit in our power accumulation function is the assumption that wealth and power are exactly convertible in the winning coalition, that is, an increase in wealth increases power in the same proportion. Jordan [17] makes a similar assumption in defining a coalition's power function that enables these coalitions to dominate others (see Section 1.3). We assume this particular function for tractability, and relegate to future research the extension to other monotonic functions of wealth and power.

[12] First, observe that in the case where agents are killed and $\phi(S^{t-1}, \pi^{t-1})$ contains more than one element, we impose no restriction in which coalition from $\phi(S^{t-1}, \pi^{t-1})$ will equal $S^t$. This allows our results to be more robust, since the evolution of the game includes any potential path of coalitions such that $S^t \in \phi(S^{t-1}, \pi^{t-1})$ for all $t$.

Finally, we assume that agents are infinitely farsighted. That is, they care about their payoff at the limit coalition formation game. This payoff, denoted by $\xi(S^\infty, \pi^\infty)$, is defined as the limit of the sequence $\xi(S^0, \pi^0), \xi(S^1, \pi^1), \xi(S^2, \pi^2)\ldots$, when it exists. If this limit does not exist, $\xi(S^\infty, \pi^\infty) = 0$.[13]

For the two sharing rules discussed above this limit exists when agents are killed. To see this, notice that there exists a time $\bar{t}$ such that $S^t$ does not change for all $t \geq \bar{t}$.[14] The convergence under PR happens because the relative power of the agents in $(S^t, \pi^t)$ will not change for any $t > \bar{t}$. The convergence under ES occurs because the agents in $S^t$ remains constant for any $t > \bar{t}$.

# 3 Agents are killed

## 3.1 Axioms

In this section we focus on the case where if a coalition $S$ forms, agents outside $S$ cannot participate in any future coalition formation process (that is, agents $i \in N \setminus S$ are "killed"). We are interested in finding transition correspondences that map to coalitions wherein agents within this coalition do not have the incentive or the power to deviate in future rounds of the game. Agents from the coalitions chosen by the transition correspondence increase their power according to their share of their prize. We introduce an axiom where a chosen coalition should continue being chosen even after power accumulates.

**Axiom 1 (Internal Self-Enforcement (ISE))** *A transition correspondence $\phi$ is* **internally self-enforcing** *(ISE) if for any game $(X, \pi)$ and any coalition $S \in \phi(X, \pi)$, we have that $S \in \phi(S, \pi_S + I\xi(S, \pi_S))$.*

Note that $I\xi(S, \pi_S)$ is the accumulated power from the prize shared by the agents inside the coalition. Therefore, ISE requires that if coalition $S$ is picked by the transition correspondence $\phi$, the same coalition $S$ will also be picked in future periods even if agents within $S$ have accumulated power (and agents who are not part of $S$ are killed). Given a correspondence $\phi$ that satisfies ISE and the game $(N, \pi)$, we say that a coalition $S$ is an *ISE-coalition* if $S \in \phi(S, \pi_S + I\xi(S, \pi_S))$.

**Axiom 2 (Dynamic Internal Rationality (DIR))** *The transition correspondence $\phi$ meets* ***dynamic internal rationality*** *(DIR) if for any game $(X, \pi)$, for any $T \in$*

[13]As we will see below, the axiom of self-enforcement guarantees that this limit always exists whether agents are killed or survive.

[14]Indeed, note that $S^0 \supseteq S^1 \supseteq S^2 \supseteq S^3 \supseteq \ldots$, thus the sequence $S^0, S^1, S^2, \ldots$ converges in finite time.

*$\phi(X,\pi)$ and for any $Z \subset X$ such that $Z \in W_{(X,\pi)}$ and $Z \in \phi(Z, \pi_Z + I\xi(Z,\pi_Z))$, we have that $Z \notin \phi(X,\pi) \Leftrightarrow \xi_i(T,\pi_T^\infty) > \xi_i(Z,\pi_Z^\infty)\ \forall i \in T \cap Z$.*

DIR requires that if there is another coalition $Z$ that is winning and internally self-enforcing in the game, then coalition $Z$ will not be chosen by the transition correspondence if and only if the agents in the intersection of a coalition $T$ that is chosen by the transition correspondence and coalition $Z$ receive a lower share of the prize in the limit by being a member of coalition $Z$.

DIR ensures that the coalition picked by the transition correspondence will yield the highest payoff for all of its members as the number of rounds approaches infinity.[15] This is similar to other notions of coalitional stability previously discussed in the literature, where a coalition is chosen if it cannot be blocked by another coalition that is winning and self-enforcing.

**Axiom 3 (Scale Invariance (SI))** *The transition correspondence $\phi$ is **scale invariant** (SI) in the vector of power if for any game $(N,\pi)$, and any coalition $S \in \phi(N,\pi)$, we have that $S \in \phi(N,\gamma\pi)\ \forall\gamma > 0$.*

Scale invariance requires that when the relative power of the agents does not change, then the coalition chosen by the transition correspondence should not change. This is a standard axiom in the literature and basically implies that we can normalize the aggregate power of the agents to 1. As we will see below, the combination of SI, with ISE and DIR, has important implications for the stability of coalitions, in particular, when varying the size of the prize.

## 3.2 Results

### 3.2.1 Proportional sharing

The class of transition correspondences that satisfy ISE, DIR and SI depends on the sharing rule. Proportional sharing will always induce an internally self-enforcing coalition because once an ISE-coalition $S^*$ forms at the initial stage, the relative power of agent $i \in S^*$ is unaffected by adding the share of the prize $\frac{\pi_i}{\pi(S^*)} \cdot I$.

Let the **minimal-power ISE transition correspondence** $\phi^*$ be defined as:

[15] DIR is weaker than the requirement that at every round the coalition chosen be the most preferred among ISE-coalitions at that time as opposed to our axiom that only requires for the coalition to be preferred at the limit. The results of this section under equal or proportional sharing will not change whether we use this alternative definition of rationality.

$$\phi^*(S,\pi) = \underset{M \in Q(S,\pi)\cup\{S\}}{\arg\min} \pi(M) \quad (1)$$

where $Q(S,\pi) = \{T \subsetneq S \mid T \in W_{(S,\pi)}, T \in \phi^*(T,\pi_T)\}$

This transition correspondence defines for the game $(S,\pi)$ a set $Q(S,\pi)$ of proper subcoalitions, which are both winning in $S$ and ISE-coalitions. It picks all coalitions that have the smallest power in $Q(S,\pi)\cup\{S\}$. If $Q(S,\pi)$ is empty then it picks coalition $S$ itself. Thus, $\phi^*(S,\pi_S) \neq \emptyset$ for any game $(S,\pi)$, hence $\phi^*$ is well-defined.

The process by which $\phi^*$ chooses a coalition can be described inductively. We illustrate this process below.

**Example 1 ($\phi^*$ and proportional sharing)** *Consider the game*

$$(S,\pi) = (\{1,2,3,4,5\},[16,17,18,20,29])$$

*and prize $I = 10$. At the outset, note that with proportional sharing, the relative power within the winning coalitions will be the same as the case when powers have not increased. For instance, if winning coalition $T = \{1,2,3\}$ forms, then agent 1's relative power is still $\frac{16}{51}$ even when the power has accumulated by distributing the prize within this coalition.*

*The inductive process by which we choose a coalition with $\phi^*$ involves the following steps. First, start from any coalition of size 1. Note that all coalitions of this size will not be chosen by $\phi^*$ since none of them are winning. The same is true for all coalitions of size 2. All coalitions of size 3 are winning. One requirement of $\phi^*$ is that the game generated by the winning coalition should map into the same coalition when other non-winning agents are killed. This requirement is satisfied by any size 3 coalition. To see this, take for example the game $(T,\pi_T) = (\{1,2,3\},[16,17,18])$ generated by the winning coalition $T$ within the game $(S,\pi)$. Coalition $T$ is winning in $S$ since $\pi(T) = 51 > \pi(S \setminus T) = 49$. Now within $T$, no coalition of size 1 would deviate since no agent has sufficient power to be winning within the game $(T,\pi_T)$. Furthermore, no coalition of size 2 is internally self-enforcing. That is, a coalition of size 2 will not deviate from $T$ since after the non-winning agent is killed, the agent with the higher power can kill the lower-powered agent in the next round. For instance if coalition $V = \{1,2\}$ deviates from $T$, then in the next round after power has accumulated, agent 2 can kill agent 1 since he has the higher relative power ($\frac{17}{33}$). Thus, agent 1 (who is farsighted) will never agree to form the coalition $V$. In this case, any coalition of* *size 3 is internally self-enforcing. Coalitions of size 4 and size 5, even though they are winning, will not be self-enforcing. This is because a coalition of size 3 will be winning in the games generated by coalitions of size 4 or 5, and those size 3 coalitions will be self-enforcing as we have shown earlier.*

*Finally, among all size 3 self-enforcing coalitions, the transition correspondence* $\phi^*$ *will choose only the coalition* $\{1, 2, 3\}$. *This is because of all size 3 coalitions that agents 1,2, or 3 are a part of, it is the coalition* $\{1, 2, 3\}$ *which maximizes their share of the prize since they have the highest relative shares in this coalition. Thus,* $T = \{1, 2, 3\} \in \phi^*(S, \pi)$.

*In contrast with the game above, for the game*

$$(M, \pi) = (\{1, 2, 3, 4, 5, 6, 7\}, [11.5, 12.5, 13, 14, 15, 16, 18])$$

*the coalition chosen by* $\phi^*$ *will be the grand coalition* $\{1, 2, 3, 4, 5, 6, 7\}$. *This is because no coalition of size 1,2 or 3 is winning within* $(M, \pi)$. *In addition, a coalition of size 3 will deviate from any coalition of size 4, 5 or 6 and will be internally self-enforcing after the non-winning agents are killed. Thus, only the grand coalition will be internally self-enforcing within the game* $(M, \pi)$.

Note that this transition correspondence is exactly the mapping that AES finds in the case where power does not accumulate. The reason this transition correspondence works in our environment where power accumulates and with proportional sharing is that, as stated earlier, the relative power of the agents in the winning coalition will not be affected by adding the prize in the same proportion as their power. Essentially, their relative power remains that the same as if the agents power were not accumulating. In all following scenarios, however, this will not be the case.

**Proposition 1** *Under proportional sharing, a transition correspondence* $\phi$ *satisfies ISE, DIR and SI if and only if* $\phi$ *is the minimal-power ISE transition correspondence.*

The proof of Proposition 1 is in the Appendix. When agents are killed, this proposition tells us that for any given game $(S, \pi)$, we can always find a unique transition correspondence that satisfies the three axioms simultaneously. By picking coalitions in the set $Q(S, \pi) \cup \{S\}$ assures us that this unique transition correspondence picks only ISE-coalitions. By picking the least-powered coalition among ISE-coalitions assures us that the agents in this coalitions will get the highest payoff (since they get a higher share of the prize) that in any ISE-coalitions to which they could possibly belong.

### 3.2.2 Equal sharing

In contrast with proportional sharing, the example below shows that with equal sharing, a transition correspondence that satisfies ISE, DIR and SI may not exist for all classes of games.

**Example 2** *[Equal sharing when power accumulates and agents are killed] Consider the game*

$$(M, \pi) = (\{1, 2, 3, 4, 5, 6, 7, 8\}, [20, 15, 14, 13, 12, 11, 10, 5])$$

*with $I = 10$ and equal sharing. Under this sharing rule, if a coalition $S$ forms within $(M, \pi)$ and continues to form forever, then the relative power of each agent at the limit approaches $\frac{1}{|S|}$. In this example, a coalition of size 3 cannot form since it does not have enough power to do so (the three highest powered agents only have power 49). A coalition of sizes 4, 5, 6 or 7 will not be internally self-enforcing since if we add the share of the prize to the agents, then a 3-person coalition can deviate and be self-enforcing, following the same logic as in Example 1. For instance, if $S = \{1, 2, 3, 4, 5, 6, 7\}$ forms, then after adding $\frac{10}{7}$ to each agent's power $T = \{1, 2, 3\}$ can deviate, since $\pi(T) = 49 + 4.3 > \pi(S \setminus T) = 46 + 5.7$. The grand coalition is not self-enforcing, since at the limit when relative powers are equalized, a 7-person coalition can deviate and be internally self-enforcing.*

A way out of this dilemma is to impose some restrictions on the class of games allowed. We say that $\tilde{G}$ is a feasible domain of games if $(S, \pi) \in \tilde{G}$ implies that $(S, \lambda\pi + \lambda\frac{I}{|S|}1_S) \in \tilde{G}$ for all $\lambda > 0$ and $I > 0$.

The restriction of the transition correspondence over a feasible domain of games is necessary for the transition correspondence that meets ISI, DIR and SI to be well-defined.[16] Thus, if we want to find a transition correspondence that meets ISI, DIR and SI, a minimal requirement on the domain of games, where the transition correspondence is defined, is for it to be feasible. This restriction is clearly not sufficient, as the entire class of games **G** is feasible, but there is no transition correspondence that meets this axiom, as illustrated in Example 2.

**Definition 3 (Strongly Balanced Game)** *A coalition formation game $(Y, \pi_Y)$ is **strongly balanced** if*

*i.* $|Y| = 2^k - 1$ *for some* $k \in \mathbb{N}$, *and*

[16]This is because if $(S, \pi) \in \tilde{G}$, then $(S, \pi + \frac{I}{|S|}1_S) \in \tilde{G}$ in order for ISE and DIR to be well-defined under ES. Moreover, in order for scale invariance to be well-defined, $(S, \lambda\pi + \lambda\frac{I}{|S|}1_S) \in \tilde{G}$ for any $\lambda > 0$.

*ii. the coalition with the $2^{k-1}$ smallest agents is a winning coalition in the game $(Y, \pi_Y)$. That is, after renaming the agents in $(Y, \pi_Y)$, if $\pi_1 \geq \cdots \geq \pi_{2^k-1}$ then*

$$\pi_1 + \cdots + \pi_{2^{k-1}-1} < \pi_{2^{k-1}} + \cdots + \pi_{2^k-1}.$$

Condition $i$ of Definition 3 restricts the cardinality of strongly balanced games to 1, 3, 7, 15, 31, etc. Condition $ii$ of Definition 3 basically requires that agents in the society have powers that are not too "far off" from each other. This restriction ensures that there will be no subset of agents from the original forming coalition that will be powerful enough to ensue a subsequent deviation among themselves. This condition is weaker than size-monotonicity assumption in the literature, where coalitions of larger size have larger power.

**Example 3 (Strongly balanced games)** *Consider the game*

$$(Y, \pi_Y) = (\{1, 2, 3, 4, 5, 6, 7\}, [24, 16, 22, 18, 26, 20, 25]).$$

*This game is strongly balanced. It clearly satisfies part i. To see that it satisfies part ii, note that the coalition with $2^{k-1}$ (in this case, 4) lowest powers is winning within $Y$, that is, $\{2, 3, 4, 6\}$ has power $\pi(\{2, 3, 4, 6\}) = 76$ while the three remaining agents $\{1, 5, 7\}$ has power $\pi(\{1, 5, 7\}) = 75$.*

*Now, consider another game*

$$(V, \pi_V) = (\{1, 2, 3, 4, 5, 6, 7\}, [12, 16, 15, 22, 13, 17, 14]).$$

*This does not satisfy Part ii of the definition since the 4 agents with the least powers are not winning within the game, that is $\pi(\{1, 3, 5, 7\}) = 54$ while $\pi(\{2, 4, 6\}) = 55$.*

**Definition 4** *The coalition formation game $(X, \pi)$ is **balanced** if it contains a coalition formation game $(Y, \pi_Y)$ such that*

*i. $Y$ is winning in $(X, \pi)$, and*

*ii. $(Y, \pi_Y + \frac{I}{|Y|} 1_Y)$ is strongly balanced.*

A balanced game contains a winning coalition such that after adding power (under equal sharing) it will become strongly balanced.

If a coalition formation game $(Y, \pi_Y)$ is strongly balanced, then it is balanced because the grand coalition $Y$ is strongly balanced even after power has been accumulated. The converse is clearly not true.

In general, for any coalition formation game $(X, \pi)$, there exists a large enough $I$ such that the game $(X, \pi)$ is balanced, since the power of the agents equalize under equal sharing by adding a large enough prize.

**Example 4 (Balanced games)** *To illustrate what a balanced game looks like, consider the game*

$$(X, \pi) = (\{1, 2, 3, 4\}, [1, 1.5, 8, 10]).$$

*First, let us take the case where $I = 1.2$ and this prize is shared equally within the coalition. It is easy to see that no coalition of size 3 generates a strongly balanced game. For instance, suppose that $\{2, 3, 4\}$ forms. After winning and splitting the prize, the power profile at round 1 is now $\pi_{\{2,3,4\}} = [1.9, 8.4, 10.4]$. Note that the agents with the two lowest powers, agent 2 and agent 3, have a combined power of 10.3 which is less than agent 4's power. Thus, with our assumptions $\{2, 3, 4\}$ does not generate a strongly balanced game and this is true for any other coalition of size 3. Therefore, Part ii of the definition of a balanced game is not satisfied.*

*However if we increase the prize sufficiently, we can show that $(X, \pi)$ is a balanced game. Consider the case where we double the prize to $I = 2.4$ and suppose that coalition $\{2, 3, 4\}$ forms. After splitting the prize equally, the new power profile for this coalition at round 1 is $\pi_{\{2,3,4\}} = [2.3, 8.8, 10.8]$. Agents 2 and 3—the two agents with the lowest powers—has a total power of 11.1 which is higher than agent 4's power. Thus, with the new prize, coalition $\{2, 3, 4\}$ generates a strongly balanced game. Therefore, the game $(X, \pi)$ is balanced since it satisfies Part ii of the definition of a balanced game.*

We say a game $(X, \pi)$ is *generic* if the power profile $\pi$ has no ties in the power of any two coalitions; that is, $\pi(S) \neq \pi(T)$ for any $S, T \subset X$. Note that the class of non-generic games has a Lebesgue measure equal to zero, so this is a weak condition.

The next result finds the largest class of games (within the set of generic games) where a transition correspondence that meets ISE, DIR and SI exists. It also finds the unique transition correspondence that meets these axioms.

**Proposition 2** *Consider a feasible domain of generic games $\tilde{G}$ and a transition correspondence $\tilde{\phi} : \tilde{G} \to 2^N$. Under equal sharing, if the transition correspondence $\tilde{\phi}$ satisfies ISE, DIR and SI, then:*

*i.* *$\tilde{G}$ should only contain balanced games.*

*ii.* $\tilde{\phi}(S, \pi) = \underset{M \in \tilde{Q}(S,\pi)}{\arg\min} |M|$

*where* $\tilde{Q}(S, \pi) = \{T | \ T \in W_{(S,\pi)} \text{ and } (T, \pi_T + \frac{I}{|T|} 1_T) \text{ is strongly balanced}\}.$

The proof of this result is in the appendix. We call $\tilde{\phi}$ the **minimal-size strongly-balanced transition correspondence**. Note that the minimal-size strongly-balanced transition correspondence $\tilde{\phi}$ picks all coalitions of the smallest size that are strongly balanced. In particular, this transition correspondence only picks coalition of size $2^k - 1$ for some $k \in \mathbb{N}$. The minimal-size strongly-balanced transition correspondence is well-defined because $\tilde{G}$ only contain balance games, thus by definition, at every game there exists at least one winning subcoalition $Y$ such that the game $(Y, \pi_Y + \frac{I}{|Y|} 1_Y)$ is strongly balanced.

The intuition behind the proof of part *ii* is that as power accumulates the relative power of the winning agents equalizes and becomes substantially larger than the losing agents. We show inductively that among vectors where power is relatively equalized, the only ISE-coalitions are of size $2^k - 1$ for some $k \in \mathbb{N}$. Therefore, an ISE-coalition must be of size $2^k - 1$ for some $k \in \mathbb{N}$, otherwise after enough rounds of power accumulation a winning subcoalition of size $2^k - 1$ for some $k \in \mathbb{N}$ will be able to deviate and be ISE. Thus, a coalition of size 5, for instance, will never be picked by the minimal-size strongly-balanced transition correspondence, since as powers equalize for these five agents after many rounds, a subcoaliton composed of any three agents can deviate and be self-enforcing.

# 4 Agents survive

## 4.1 Motivating examples

Let us begin this section with two examples to motivate why both the minimal-power ISE transition correspondence $\phi^*$ and the minimal-size strongly-balanced transition correspondence $\tilde{\phi}$ introduced in the previous section do not work when agents survive.

**Example 5 (Equal sharing and $\tilde{\phi}$)** *Consider the game*

$$(S, \pi) = (\{1, 2, 3, 4, 5\}, [23, 21, 20, 19, 17])$$

*with prize $I = 1$. We look at the case of equal sharing and the minimal-size strongly-balanced transition correspondence $\tilde{\phi}$. To contrast with the previous section, first note that this game is balanced. Suppose we proceed with the transition correspondence $\tilde{\phi}$. Thus, the coalition $\{3, 4, 5\}$ is an element of $\tilde{\phi}$. However, at the $67^{th}$ round the powers of agents 3 and 4 who originally had powers 20 and 19, respectively, now have powers 42.33 and 41.33. Thus, the combined powers of agents 3 and 4 is higher in that round than the rest of society composed of Agents 1,2 and 5 (since agents 1 and 2 still survive),* *that is,* $\pi^{67}_{\{3,4\}} = 83.66 > \pi^{67}_{\{1,2,5\}} = 83.33$. *Thus,* $\{3,4\}$ *can deviate from* $\{3,4,5\}$ *and be self-enforcing since neither 3 nor 4 have sufficient power to win, even if power has accumulated. Hence, the original coalition* $\{3,4,5\}$ *is not self-enforcing.*

*The coalition* $\{2,3,4,5\}$*, however, will be self-enforcing since no member of this coalition will want to deviate in any round. To see this, observe that any coalition of size 3 will not deviate from* $\{2,3,4,5\}$ *since in some future round a coalition of size 2 will have enough relative power to deviate from the 3-agent coalition and be self-enforcing. To illustrate, suppose that* $\{3,4,5\}$ *deviates from* $\{2,3,4,5\}$ *at round 2. After sharing the prize equally among the three of them starting from round 2, at round 68 agents 3 and 4 can deviate since* $\pi^{68}_{\{3,4\}} = 84.17 > \pi^{68}_{\{1,2,5\}} = 83.83$ *and this new 2 person coalition will be self-enforcing (since neither 3 nor 4 can have higher individual power than the rest of society). Moreover, any size 2 coalition cannot deviate from* $\{2,3,4,5\}$ *since in no succeeding round will 2 agents have enough combined power to dominate the rest of the society.*

*As we will show below, a self-enforcing coalition under equal-sharing must have size* $2^k$ *for some* $k \in \mathbb{N}$*, unlike the minimal-size strongly-balanced transition correspondence* $\tilde{\phi}$ *which chooses coalitions of size* $2^k - 1$ *for some* $k \in \mathbb{N}$.

**Example 6 (Proportional sharing and** $\phi^*$**)** *For the same game as in example 5 above, we look at the case of proportional sharing and the minimum-power self-enforcing transition correspondence* $\phi^*$.

*With* $\phi^*$*, coalition* $\{3,4,5\}$ *forms. At the third round after power accumulates, agent 3 (*$\pi^3_3 = 21.07$*) has now a higher power than agent 2 (*$\pi^3_2 = 21$*). Therefore agents 4 and 5 will get a higher proportion of the prize if they align with agent 2, and therefore* $\{3,4,5\}$ *will deviate to* $\{2,4,5\}$ *at the third round. The original coalition generated by* $\phi^*$ *is therefore not self-enforcing. This phenomenon of jumping from one coalition to another always occur for proportional sharing and* $\phi^*$ *when power accumulates. In order to avoid it, we introduce below a class of priority transition correspondences.*

## 4.2 Axioms

We introduce the analogous version of ISE for the case where agents survive throughout rounds. Similarly to ISE, once a coalition is chosen, it must continue to be chosen even after power accumulates and agents survive.

**Axiom 4 (External Self-Enforcement (ESE))** *A transition correspondence* $\phi$ *is* **externally self-enforcing** *(ESE) if for any game* $(N, \pi)$ *and* $S \in \phi(N, \pi)$*, we have that* $S \in \phi(N, \pi + I\xi^N(S, \pi))$.

When there is no confusion, we say that a coalition $S$ is an *ESE-coalition* if it is generated by the transition correspondence satisfying ESE. This modification shows that coalitions that are externally self-enforcing should map into the same coalition even though agents from $N \setminus S$ (agents outside $S$) can still form coalitions and threaten $S$.

We also introduce an analogous version of rationality in this section where agents survive. Dynamic external rationality (DER) requires that the transition correspondence will pick among all ESE-coalitions the coalition that gives the highest payoff to the agents at the limit.[17]

**Axiom 5 (Dynamic External Rationality (DER))** *A transition correspondence $\phi$ meets* ***dynamic external rationality*** *(DER) if for any $T \in \phi(N, \pi)$ and for any $Z \subset N$ such that $Z \in W_{(N,\pi)}$ and $Z \in \phi(N, \pi + I\xi^N(Z, \pi))$, we have that $Z \notin \phi(N, \pi) \Leftrightarrow \xi_i(N, \pi_T^\infty) > \xi_i(N, \pi_Z^\infty)\ \forall i \in T \cap Z$.*

There are many transition correspondences that satisfy ESE, DER and SI. The trivial transition correspondence $\phi(N, \pi) = N$ for any $\pi$ is one of them. In order to avoid these trivial correspondences, we also introduce the axiom of *independence of zeros*, which says that the agents without power should not affect the winning coalition.

**Axiom 6** *A transition correspondence $\phi$ is* ***independent of zeros (IZ)*** *whenever $\phi(N, [\pi_S, 0_{N \setminus S}]) = \phi(S, \pi_S)$.*

In particular, notice that IZ implies that if an agent has no power, he will not be chosen by the correspondence. IZ, along with the continuity of the transition correspondence, allows us to compare games where $k$ agents are participating (i.e. exactly $k$ agents have non-zero power) with games where $\tilde{k}$ agents are participating, by making the power of $|k - \tilde{k}|$ agents tend to zero. Thus, it is related to the consistency axiom in other settings such as resource allocation problems (Thomson [37, 38]). Section 4.3.1 introduces transition correspondences that do not satisfy IZ.

## 4.3 Results

We find that under equal sharing, there is always a transition correspondence that satisfies the axioms above. However, this is not true under proportional sharing.

To formalize the results, define the **minimal $2^k$-sized transition correspondence** $\phi^{**}$ as follows:

[17] Note this limit exists by ESE because the coalition that is chosen at time 0 is also chosen at any time in the future.

$$\phi^{**}(N,\pi) = \arg \min_{M \in Q(N,\pi)} |M| \tag{2}$$

where $Q(N,\pi) = \{S \in 2^N$ such that $S \in W_{(N,\pi)}$ and $|S| = 2^m$ for some $m \in \mathbb{N}\}$.

The minimal $2^k$-sized transition correspondence $\phi^{**}(N,\pi)$ chooses all least-sized winning coalitions that are of size $1, 2, 4, 8, 16, \ldots$ This transition correspondence is well-defined since there always exists a winning coalition of size $2^k$ for some $k \in \mathbb{N}$ and therefore $\phi^{**}(N,\pi) \neq \emptyset$.

**Proposition 3** *i. Under equal sharing, the correspondence $\phi^{**}$ is the only transition correspondence that satisfies ESE, DER, SI and IZ.*

*ii. Under proportional sharing, there is no transition correspondence that satisfies ESE, DER, SI, and IZ.*

The proof of this result is in the appendix. The intuition behind the proof of part *i* is that as power accumulates the relative power of the winning agents equalizes and becomes substantially larger than the losing agents. We show inductively that among vectors where power is relatively equalized, the only ESE-coalitions are of size $2^k$ for some $k \in \mathbb{N}$. Therefore, for a general power vector, an ESE-coalition must be of size $2^k$ for some $k \in \mathbb{N}$, otherwise after enough rounds of power accumulation a winning subcoalition of size $2^k$ for some $k \in \mathbb{N}$ will be able to deviate and be ESE. For instance, if 8 agents form and continue to form after many rounds, no agents of size 4 can deviate since the other 4 sidelined agents can partner with some of the other surviving non-winning agents to overthrow the 4 that deviated. They can do that since the 8 agents have relatively equal power.

The proof of part *ii* shows that for games with three or more agents the axioms are incompatible. The key argument relies on the fact that for coalition formation games with a dictator (an agent who has enough power to win by himself) the dictator must be chosen. Therefore, ESE implies that in games for three agents without a dictator a coalition of size 2 or less cannot be chosen, otherwise as power accumulates a game with a dictator will form. Finally, choosing the grand coalition in games for three agents without a dictator will contradict IZ.

### 4.3.1 Proportional sharing without IZ

In this subsection, we study transition correspondences that satisfy ESE, DER and SI but do not necessarily satisfy IZ.

**Definition 5** *A* ***feasible sequence of coalitions*** *is a finite sequence of coalitions from* $N$*, denoted* $\{S^0, S^1, \ldots, S^k\}$*, such that*

*i.* $S^k = N$

*ii. if* $S^i \cap S^j \neq \emptyset$ *and* $i < j$ *then* $S^i \subsetneq S^j$

*Denote the set of all feasible sequences of coalitions as* $\mathcal{F}$*.*

**Definition 6** *Given a feasible sequence of coalitions* $\{S^0, S^1, \ldots, S^k\} \in \mathcal{F}$*, we define a* ***sequential transition correspondence*** *as:*

$$\bar{\phi}(N, \pi) = \arg \min_{M \in Q \cup \{N\}} \pi(M)$$

*where* $Q(N, \pi) = \{S^i \mid S^i \in W_{(N,\pi)}, \text{ such that for all } S^j \subset S^i,\ S^j \notin W_{(S^i, \pi_{S^i})}\}$*.*

In any feasible sequence of coalitions, the sequential transition correspondence will pick the coalition with the least power that is winning and such that it does not contain a subset that is winning within that coalition. If that set is empty, then the grand coalition is picked. Thus, given any feasible sequence of coalitions, $\bar{\phi}(N, \pi)$ always picks either the grand coalition or a subcoalition of it. Hence, $\bar{\phi}(N, \pi) \neq \emptyset$ and is well-defined.

**Example 7** *Suppose there are seven agents* $\{1, 2, 3, 4, 5, 6, 7\}$ *a feasible sequence of coalitions given by*

$$\{4\}, \{3, 4, 5\}, \{1, 2, 3, 4, 5\}, \{1, 2, 3, 4, 5, 6, 7\}.$$

*Consider the power profile* $\pi = [12, 12.5, 13, 13.5, 14, 15, 16]$*. The winning coalition with the least power that appears in the sequence is coalition* $\{1, 2, 3, 4, 5\}$*. However, the sequential transition correspondence will not pick this coalition, since coalition* $\{3, 4, 5\}$ *is winning within* $\{1, 2, 3, 4, 5\}$*. The sequential transition correspondence then picks the grand coalition in this case.*

*Now consider the power profile* $\pi' = [10.5, 12, 14, 18, 19, 13, 13.5]$*. The winning coalition with the least power is* $\{3, 4, 5\}$ *and this will be picked by the sequential transition correspondence since* $\{4\}$ *is not winning within* $\{3, 4, 5\}$*.*

Intuitively, the restriction to the set of feasible sequence of coalitions precludes the ability of agents to include lower-powered agents to increase their own payoffs. By limiting the possible deviations only into a subcoalitions (given by the sequence), there can be no “jumping”.

**Proposition 4** *Under proportional sharing, any sequential transition correspondence satisfies ESE, DER and SI.*

The intuition behind the proof of this proposition is as follows. As power accumulates, the relative power of the winning agents remains the same while the relative power of the non-winning agents tends to zero. However, by construction of the correspondence, no winning subcoalition from the winning coalition is feasible. Hence, the winning coalition is chosen throughout time even after power accumulates.

# 5 Conclusion

This paper develops an axiomatic approach to a coalition formation model by focusing on two main axioms: self-enforcement and rationality. Our results highlight that despite the simplicity of our model's environment, there are still cases where finding rules on choosing coalitions that satisfy these two axioms is not straightforward. By restricting some of our game's parameters, we have provided rich characterizations of possible transition correspondences that do satisfy these axioms. This is summarized in the table below.

| Scenario | Equal Sharing | Proportional Sharing |
|---|---|---|
| Agents Killed | minimal-size strongly-balanced TC $\tilde{\phi}$<br>for balanced games<br>(See Proposition 2(ii)) | minimal-power ISE TC $\phi^*$<br>(See Equation 1) |
| Agents Survive | minimal $2^k$-sized TC $\phi^{**}$<br>(See Equation 2) | sequential TC $\bar{\phi}$<br>for a feasible sequence of coalitions<br>(See Definition 6) |

In the case where agents are killed and the resource is distributed equally among agents in the winning coalition, a self-enforcing and rational transition correspondence exists (the *minimal-size strongly-balanced transition correspondence*) only if we can find a subset of the grand coalition of size $2^k - 1$ (i.e., of size 1,3,7,15, 31,...) that is strongly balanced after power has been added (Section 3.2, Proposition 2). Under the same case but with proportional sharing of the resource, the axioms of self-enforcement and rationality uniquely determines the *minimal-power ISE transition correspondence*, which picks the coalition(s) of the smallest power among self-enforcing coalition (Section 3.2, Proposition 1). When agents survive, under proportional sharing we have to restrict to feasible sequences of partitions in order to find a self-enforcing and rational transition correspondence, which we call the *sequential transition correspondence*

(Section 4.3.1, Proposition 4). Under equal sharing the *minimal* $2^k$*-sized transition correspondence* picks the smallest coalition of size $2^k$ (i.e., size, 1,2,4,8,16, ...) that is winning (Section 4.3, Proposition 3).

We believe that the attractiveness of our model's diverse results derives from highlighting the observation that the different institutions, or "rules of the game" (e.g., power, sharing rules, feasible coalitions) surrounding the coalition formation process play a large role in determining the types of coalitions that will be self-enforcing through time.

## 5.1 Extensions and open questions

The results in this paper can be easily extended to the case of supermajority, that is, when the set of winning coalitions is defined as $W_{(T,\pi)} = \{S \subset T | \pi(S) > \alpha\pi(T \setminus S)\}$ for $\alpha \in [0.5, 1)$. For instance, consider the case of $\phi^*$. Under supermajority, $\phi^*$ picks the coalition with the highest share among the winning and self-enforcing coalitions.

The generalization of $\bar{\phi}$ is also straightforward. It picks the winning coalition with the least power in a feasible sequence without any subcoalition that is winning. For instance, in Example 7 with power profile $\pi'$, if $\alpha = .70$ then the coalition $\{1,2,3,4,5\}$ will be the winning coalition with the least power in the feasible sequence. This will be picked by the transition correspondence $\bar{\phi}$ since the coalition $\{3,4,5\}$ is not winning within $\{1,2,3,4,5\}$ because it has less than 70% of the power.

The generalization of $\phi^{**}$ is more involved but still straightforward. The first step is to create for a given $\alpha \in [0.5, 1)$ a sequence of admissible sizes $A^\alpha$ that follows the next inductive process. The first element of the sequence is $A^\alpha_1 = 1$ since a coalition of size 1 is (externally) self-enforcing. For any $k > 1$, $A^\alpha_k$ is the smallest integer such that $A^\alpha_k > A^\alpha_{k-1}$ and $\frac{A^\alpha_{k-1}}{A^\alpha_k} \leq \alpha$.

For instance, for $\alpha = .5$ the set of admissible sizes is $\{1,2,4,8,16,32,\dots\}$. For $\alpha = .75$ the set of admissible sizes is $\{1,2,3,4,6,8,\dots\}$. For $\alpha = .6$ the set of admissible sizes is $\{1,2,4,7,12,20,\dots\}$.

The mapping $\phi^{**}$ is therefore modified as:

$$\phi^{**}(N,\pi) = \arg \min_{M \in Q \cup \{N\}} |M|$$

where $Q = \{S \in 2^N$ such that $S \in W_{(N,\pi)}$ and $|S| \in A^\alpha\}$.

There are multiple open questions for future research. A first extension of our study is the implementation of the stable coalitions found by this paper as the equilibrium of a non-cooperative game (see Acemoglu et al. [1]). Second, there is a need to develop a new theory describing how the different sharing rules emerge endogenously from the coalition formation process. Third, the case where the decision to kill agents is endogenous (as well as the possibility of growth in the number of agents in the society) should also be examined. Fourth, the extension to more general functions of power accumulation should be studied. Finally, another extension to study is the role of externalities (such as cultural characteristics or religion) in the coalition formation process (see Juarez [18, 19] for some advances).

# 6 Appendix

## 6.1 Proofs of Section 3

***Proof of Proposition 1***

We begin the proof by introducing two axioms which are variants of ISE and DIR.

**Axiom 7 (Self-enforcement (SE))** *The transition correspondence $\phi$ is* ***self-enforcing (SE)*** *if for any game $(X, \pi) \in \mathbf{G}$ and $S \in \phi(X, \pi)$, then $S \in \phi(S, \pi_S)$.*

**Axiom 8 (Rationality (RAT))** *The transition correspondence $\phi$ is* ***rational (RAT)*** *if for any $S \in 2^N$, for any $T \in \phi(S, \pi)$ and for any $Z \subset S$ such that $Z \in W_{(S,\pi)}$ and $Z \in \phi(Z, \pi_Z)$, we have that $Z \notin \phi(S, \pi) \Leftrightarrow \xi_i(T, \pi_T) > \xi_i(Z, \pi_Z)\ \forall i \in T \cap Z$.*

We will show in subsequent steps that ISE and SI imply the axiom SE above and DIR and SI imply RAT. We then show that $\phi^*$ is the unique transition correspondence that satisfy SE and RAT.

**Proof.** Step 1. Under proportional sharing, ISE and SI imply SE.

Proof. If $S \in \phi(X, \pi)$, then by ISE, $S \in \phi(S, \pi_S + I \cdot PR(S, \pi_S))$. Since the shares in $PR(S, \pi_S)$ are split in the ratio of $\pi_S$, then by scale invariance, $S \in \phi(S, \pi_S)$.

Step 2. Under proportional sharing, DIR and SI imply RAT.

Proof. Consider any $X \in 2^N$, and subset $T \in \phi(X, \pi)$ and any $Z \subset X$ such that $Z \in W_{(X,\pi)}$ and $Z \in \phi(Z, \pi_Z + I \cdot PR(Z, \pi_Z))$. Then, $Z \in \phi(Z, \pi_Z)$ by Step 1.

Therefore, we have that

$$Z \notin \phi(X, \pi) \Leftrightarrow \xi_i(T, \pi_T^{\infty}) > \xi_i(Z, \pi_Z^{\infty})\ \forall i \in T \cap Z$$

$$\Leftrightarrow \xi_i(T, \pi_T) > \xi_i(Z, \pi_Z)\ \forall i \in T \cap Z$$

because the sharing rule is proportional. Hence, RAT is satisfied.

Step 3. There exists a unique transition correspondence that meets SE and RAT. That transition correspondence is $\phi^*$.

The proof of this step is similar to Theorem 1 of Acemoglu, et al [1] (AES) and will not be shown here. They prove this for the case where power does not accumulate, agents are killed and the sharing rule is proportional. Steps 1 and 2 in this proof basically transforms the axioms ISE and DIR to Acemoglu et al's axioms SE and RAT. Jandoc and Juarez [16] extends AES results to a more general class of sharing rules that satisfy the property of "consistent ranking" where agents have the same ordinal ranking over coalitions in which they belong. Equal and proportional sharing meet consistent ranking.

Step 4. We show that $\phi^*$ satisfies SI, ISE and DIR

- $\phi^*$ satisfies SI.

  Since for any $\gamma > 0$

$$\phi^*(S, \gamma\pi) = \underset{M \in Q(S,\pi)\cup\{S\}}{\arg\min}\ \gamma\pi(M) = \underset{M \in Q(S,\pi)\cup\{S\}}{\arg\min}\ \pi(M) = \phi^*(S, \pi)$$

- $\phi^*$ satisfies ISE.

  For the game $(S, \pi)$ take a coalition $V \in \phi^*(S, \pi)$. By the definition of $\phi^*$, either $V = S$ or $V \in Q(S, \pi)$. If $V = S$ then $S \in \phi^*(S, \pi)$. By SI, we have that $S \in \phi^*(S, \pi(1 + \frac{I}{\pi(S)})) = \phi^*(S, \pi + I \cdot PR(S, \pi_S))$.

  On the other hand, if $V \in Q(S, \pi)$, then $V \in \phi^*(V, \pi_V)$ by definition in Equation 1. By SI, we have that $V \in \phi^*(V, \pi_V(1 + \frac{I}{\pi(V)})) = \phi^*(V, \pi_V + I \cdot PR(V, \pi_V))$.

- $\phi^*$ satisfies DIR.

  Consider the game $(S, \pi)$. Let $V \in \phi^*(S, \pi)$ and $Y \subset S$ such that $Y \in W_{(S,\pi)}$ and $Y \in \phi^*(Y, \pi_Y + I \cdot PR(Y, \pi_Y))$.

  First, consider the case where $Y \notin \phi^*(S, \pi)$. By definition,

$$V \in \underset{M \in Q(S,\pi)\cup\{S\}}{\arg\min}\ \pi(M).$$

  Therefore, $\pi(Y) > \pi(V)$. Therefore, for every agent $i \in Y \cap V$, $PR_i(V, \pi_V) > PR_i(Y, \pi_Y)$. Thus, $\xi_i(V, \pi_V^\infty) > \xi_i(Y, \pi_Y^\infty)$.

On the other hand, if $\xi_i(V, \pi_V^\infty) > \xi_i(Y, \pi_Y^\infty)$ for all $i \in V \cap Y$, then, $\pi(Y) > \pi(V)$. Therefore,

$$Y \notin \underset{M \in Q(S,\pi) \cup \{S\}}{\arg\min} \pi(M).$$

Hence, $Y \notin \phi^*(S, \pi)$.

■

***Proof of Proposition 2***

**Proof.** We prove four intermediate steps before proving part $i$ and $ii$ in step 5.

For all the steps, we consider a transition correspondence $\phi$ that satisfies ISE, DIR and SI.

Step 1: If a coalition $S$ is picked by any transition correspondence $\phi$ and continues to form, then over time the relative power of $i \in S$ approaches $\frac{1}{|S|}$.

Proof. Note that under equal sharing, the relative power of agent $i$ in a coalition $S$ that continues to form through the $k^{th}$ stage is $\frac{\pi_i^0 + \frac{kI}{|S|}}{\sum_{i \in S} \pi_i^0 + kI}$. Evaluating this expression as $k \to \infty$ by using l'Hopital's rule yields:
$\lim_{k\to\infty} \frac{\pi_i^0 + \frac{kI}{|S|}}{\sum_{i \in S} \pi_i^0 + kI} = \lim_{k\to\infty} \frac{\frac{I}{|S|}}{I} = \lim_{k\to\infty} \frac{1}{|S|} = \frac{1}{|S|}$

Step 2: Any coalition that is chosen by a transition correspondence $\phi$ that satisfies ISE and DIR should be of size $2^m - 1$ for some $m \in \mathbb{N}$.

Proof. Consider any coalition $S$ that is chosen by the transition correspondence $\phi$, and suppose that $|S| = 2^m - 1 + r$ and $r \in [0, 2^m - 1]$. We will prove this step by induction on $m$.

Consider the base of induction $m = 1$. In this case,

$$|S| = \begin{cases} 1 \text{ if } r = 0 \\ 2 \text{ if } r > 0 \end{cases}$$

We know that $|S| = 1$ is an ISE-coalition since a singleton maps into itself. On the other hand, if $|S| = 2$, then the agent $i$ such that $\pi_i > \pi_j$ can always deviate from $S$ and be self-enforcing (since he is a singleton coalition). Thus, $S$ where $|S| = 2$ is not an ISE-coalition.

Let our induction hypothesis be that the statement is true for $m = h$. That is,

$$\text{If } |S| = 2^h - 1 + r \text{ then } \begin{cases} S \text{ is an ISE coalition if } r = 0 \\ S \text{ is not an ISE coalition if } r > 0 \end{cases}$$

We now show that this relationship remains true for $m = h + 1$.

If $r = 0$, then:

- By Step 1 the relative power of $i \in S$ is $\frac{1}{2^{h+1}-1}$ as the rounds approach infinity. That is, $\lim_{k\to\infty} \frac{\pi^k}{\pi^k(S)} = \left[\frac{1}{2^{h+1}-1}, \frac{1}{2^{h+1}-1}, \ldots, \frac{1}{2^{h+1}-1}\right]$
- A coalition $T$ that wishes to deviate from $S$ must be at least $2^h - 1 + r$, where $2^h - 1 \leq 2^h - 1 + r < 2^{h+1} - 1$. Note that a $|T| = 2^h - 1$ will not be winning since $\pi(N \setminus T) > \pi(T)$
- In this case, by Step 1 we know that if $T$ continues to form, then the relative power of $i \in T$ will approach $\frac{1}{2^h-1+r}$ in the limit.
- By the same reasoning, a coalition $V$, where $|V| = 2^h - 1$ can deviate from coalition $T$. This will be an ISE-coalition by our induction hypothesis. Thus $T$ is not an ISE-coalition. Therefore, $S$ where $|S| = 2^{h+1} - 1$ is an ISE-coalition.

If $r > 0$, then:

- By Step 1 the relative power of $i \in S$ is $\frac{1}{2^{h+1}-1+r}$ as the rounds approach infinity. That is, $\lim_{k\to\infty} \frac{\pi^k}{\pi^k(S)} = \left[\frac{1}{2^{h+1}-1+r}, \frac{1}{2^{h+1}-1+r}, \ldots, \frac{1}{2^{h+1}-1+r}\right]$
- A coalition $T$ where $|T| = 2^h - 1$ can deviate from $S$. From our induction hypothesis $T$ will be an ISE-coalition. Therefore $S$ where $|S| = 2^h - 1 + r$ cannot be an ISE-coalition if $r > 0$.

Step 3: Consider the transition correspondence $\phi$ that satisfies ISE and DIR and the coalition formation game $(Y, \pi_Y)$ that is strongly balanced. Then, $\phi(Y, \pi_Y) = \{Y\}$.

Proof. By step 2, the coalitions chosen by $\phi$ must be of size $2^k - 1$. Since $(Y, \pi_Y)$ is strongly balanced, the only winning coalition of this type is $Y$. Hence, $\phi(Y, \pi_Y) = \{Y\}$.

Step 4: If the coalition $V$ is chosen by the transition correspondence that satisfies ISE and DIR at game $(X, \pi)$, then $(V, \pi_V + \frac{I}{|V|} 1_V)$ is strongly balanced.

Proof.

By step 2, $|V| = 2^k - 1$ for some $k \in N$. Suppose that $(V, \pi_V + \frac{I}{|V|} 1_V)$ is not strongly balanced.

Partition $V$ into the disjoint coalitions $S, T$ and $U$, that is $V = S \cup T \cup U$, and such that coalition $S$ contains the $2^{k-2} - 1$ largest elements in $V$, coalition $T$ contains the $2^{k-2}$ largest elements in $V \setminus S$ and coalition $U$ contains $V \setminus (S \cup T)$, which are the $2^{k-1}$ smallest elements in $V$.

Since $(V, \pi_V + \frac{I}{|V|} 1_V)$ is not strongly balanced then $\pi(S \cup T) > \pi(U)$.

By step 1, since power equalizes under ES, there exists a time $t$ such that $U$ becomes a winning coalition in $V$. That is,

$$\pi(S \cup T) + tI(2^{k-1} - 1) < \pi(U) + tI2^{k-1}.$$

Therefore,

$$\pi(S \cup T) - \pi(U) < tI \tag{3}$$

Also, note that by the choice of $S, T$ and $U$,

$$\pi(S) - \pi(T) \leq \pi(S \cup T) - \pi(U) \tag{4}$$

To see this, assume that $\pi(S) - \pi(T) > \pi(S \cup T) - \pi(U)$, then $\pi(T) < \frac{\pi(U)}{2}$. Since $2|T| = |U|$, then $\frac{\pi(T)}{|T|} < \frac{\pi(U)}{|U|}$, which contradicts the choice of $T$ and $U$, therefore inequality 4 holds.

Combining inequalities 3 and 4:

$$\pi(S) - \pi(T) < tI$$

Therefore,

$$\pi(S) + tI(2^{k-2} - 1) < \pi(T) + tI2^{k-2} \tag{5}$$

The left and right hand side of inequality 5 are the power of coalition $S$ and $T$, respectively, after $t$ rounds. Hence, coalition $S \cup T$ is strongly balanced at time $t$. Moreover, by the choice of $t$, coalition $S \cup T$ is winning in $V$ at time $t - 1$, hence it is balanced. Thus, $S \cup T$ can deviate from $V$ at time $t - 1$ and be internally self-enforcing by step 3. This contradicts DIR.

Step 5: Proofs of parts $i$ and $ii$.

Part $i$ follows directly from step 4, because if the transition correspondence selects $Y$ at the coalition formation game $(X, \pi)$, then $Y$ is winning in $(X, \pi)$ and $(Y, \pi_Y)$ is strongly balanced. Hence, $(X, \pi)$ is balanced.

To prove part $ii$, consider the balanced game $(X, \pi)$ and suppose that $Y \in \phi(X, \pi)$. By step 2, $|Y| = 2^m - 1$ for some $m$. Suppose that $Z \in W_{(X,\pi)}$, $(Z, \pi_Z)$ is strongly balanced and $Z \notin \phi(X, \pi)$. Since $(Z, \pi_Z)$ is strongly balanced, then $Z \in \phi(Z, \pi_Z)$ by step 3. Hence, by ISE, $Z \in \phi(Z, \pi_Z + \frac{I}{|Z|} 1_Z)$. Therefore, by DIR, $\xi_i(Y, \pi_Y^\infty) > \xi_i(Z, \pi_Z^\infty)$ for all $i \in Y \cap Z$. Hence, by step 1, $\xi_i(Y, \pi_Y^\infty) = \frac{1}{|Y|}$ and $\xi_i(Z, \pi_Z^\infty) = \frac{1}{|Z|}$. Thus, $|Y| < |Z|$.

■

## 6.2 Proofs of Section 4

***Proof of Proposition 3***

**Proof of part $i$.**

We fix the transition correspondence $\phi^{**}$ that satisfies ESE, DER, SI, IZ.

Step 1: If a coalition $S$ is chosen by the transition correspondence $\phi^{**}$ that satisfies ESE, then over time the relative power of $i \in S$ approaches $\frac{1}{|S|}$.

Proof. This is similar to the proof of Step 1 in Proposition 2, therefore it is omitted.

Step 2: Consider a coalition formation game $(N, \pi)$ such that $\phi^{**}(N, \pi) = \{S\}$. Then, there exists an open set $B \subset \mathbb{R}_+^N$ such that $\pi \in B$ and $S \in \phi^{**}(N, \tilde{\pi})$ for any $\tilde{\pi} \in B$.

Proof. Suppose that is not the case. Then, there exists a sequence of power vectors $\{\bar{\pi}^i\}_i$ where $S$ is winning in the coalition formation $(N, \bar{\pi}^i)$ for every $i$, $\lim_{i \to \infty} \bar{\pi}^i = \pi$ and $S \notin \phi^{**}(N, \bar{\pi}^i)$ for all $i$.

Since $\phi^{**}(N, \bar{\pi}^i)$ is chosen from the finite set $2^N$, then we can find a set that is chosen an infinite number of times in the sequence. That is, we can find $T \subset N$ and a subsequence $\{\tilde{\pi}^i\}_i \subset \{\bar{\pi}^i\}_i$ such that $T \in \phi^{**}(N, \tilde{\pi}^i)$ for all $i$. Hence, by continuity, $T \in \phi^{**}(N, \pi)$, which is a contradiction.

Step 3: Let $S^m = \{1, \ldots, m\}$ and consider the coalition formation game $(N, \pi^m)$ such that $\pi_j^m = 1$ if $j \in S^m$, and $\pi_l^m = 0$ if $l \in N \setminus S^m$. We will show by induction on $m$ that if $m = 2^k$ for some $k \in \mathbb{N}$ then $\phi^{**}(N, \pi^m) = \{S^m\}$; and if $m = 2^k + r$ for some $k \in \mathbb{N}$ and $0 < r \leq 2^k - 1$ then $\phi^{**}(N, \pi^m) = \{T | T \subset S^m \text{ and } |T| = 2^k\}$.

Proof. First, we start with the base of induction $m = 1$. By IZ, $\phi^{**}(N, \pi^1) = S^1$.

Second, suppose that the statement is true for $m < i$. We will show that it is also true for $m = i$.

Let $i = 2^k + s$ for $s \in [0, 2^k - 1]$ and $T \in \phi^{**}(N, \pi^i)$. By IZ, $T \subset S^i$. Since $T$ is winning in $(N, \pi^i)$, then $|T| > 2^{k-1}$. Thus, $|T| = 2^{k-1} + r$ for $0 < r \leq 2^{k-1} + s$. By step 1, scale invariance and continuity, $T \in \phi^{**}(N, \pi^T)$ where $\pi_i^T = 1$ if $i \in T$ and $\pi_i^T = 0$ if $i \in N \setminus T$. Thus, up to renaming the agents, $S^{2^{k-1}+r} \in \phi^{**}(N, \pi^{2^{k-1}+r})$. By the induction hypothesis, for $r \neq 2^{k-1}, 2^{k-1} + s$, we have that $S^{2^{k-1}+r} \notin \phi^{**}(N, \pi^{2^{k-1}+r})$. Thus, $r = 2^{k-1} + s$ or $r = 2^{k-1}$.

We analyze the next three cases depending on whether $s = 0$ or $s \neq 0$, and whether $r = 2^{k-1} + s$ or $r = 2^{k-1}$.

Case 1. Suppose $s = 0$.

Then, $r = 2^{k-1}$. Therefore $T = S^i$. Thus, $\phi^{**}(N, \pi^i) = \{S^i\}$.

Case 2.1. $s \in (0, 2^k - 1]$ and $r = 2^{k-1} + s$.

Then, $|T| = 2^{k-1} + 2^{k-1} + s = 2^k + s$. Thus, $S^{2^k+s} \in \phi^{**}(N, \pi^{2^k+s})$. Consider the vector $v^t \in \mathbb{R}_+^N$ such that

$$v^t = (tI)\pi^{2^k} + (\epsilon + \delta)\pi^{2^k+s} + (\delta)(\pi^{2^k+s} - \pi^{2^k}),$$

where $\delta < 2^k\epsilon < I$.

Note that for every $t \geq 0$, the size of the smallest winning coalition that is a power of 2 equals $2^k$.

Given $\epsilon$ and $\delta$, note that as $t$ tends to infinity, the relative power equalizes. Thus, by step 2 there exists a large $t^*$ such that $S^{2^k} \in \phi^{**}(N, v^{t^*})$.

Similarly, any winning coalition that is a power of 2 in the game $(N, v^{t^*-1})$ should be of size $2^k$. Since $S^{2^k} \in \phi^{**}(N, v^{t^*-1} + I\pi^{2^k})$, and $S^{2^k}$ is winning in $(N, v^{t^*-1})$, then by DER, $S^{2^k} \in \phi^{**}(N, v^{t^*-1})$.

Similarly, $S^{2^k} \in \phi^{**}(N, v^{t^*-2}), S^{2^k} \in \phi^{**}(N, v^{t^*-3}), \ldots$.

We can continue like that until $t = 0$. Therefore,

$$S^{2^k} \in \phi^{**}(N, (\epsilon + \delta)\pi^{2^k+s} + (\delta)(\pi^{2^k+s} - \pi^{2^k})).$$

By continuity, as $\delta \to 0$,

$$S^{2^k} \in \phi^{**}(N, \epsilon\pi^{2^k+s}).$$

By SI,

$$S^{2^k} \in \phi^{**}(N, \pi^{2^k+s}).$$

By assumption, $S^{2^k+s} \in \phi^{**}(N, \pi^{2^k+s})$, which contradicts DER, since the agents in $S^{2^k}$ strongly prefer $S^{2^k}$ to $S^{2^k+s}$.

Case 2.2. Supppose $s \in (0, 2^k - 1]$ and $r = 2^{k-1}$.

In this case, $S^{2^k} \in \phi^{**}(N, \pi^{2^k+s})$. Since all coalitions of size $2^k$ give exactly the same payoff and $\phi^{**}$ satisfies ESE, then $\phi^{**}(N, \pi^i) = \{T|T \subset S^i \text{ and } |T| = 2^k\}$.

Step 4: For any coalition formation game $(N, \pi)$, if $S \in \phi^{**}(N, \pi)$, then $|S| = 2^k$ for $k \in \mathbb{N}$.

Proof.

Let $e^S \in \mathbb{R}^N$ be the vector such that $e_i^S = 1$ if $i \in S$ and $e_i^S = 0$ if $i \in N \setminus S$.

By ESE, $S \in \phi^{**}(N, \pi + tIe^S)$ for $t = 1, 2, \ldots$.

By SI, $S \in \phi^{**}(N, \frac{\pi}{tI} + e^S)$ for $t = 1, 2, \ldots$.

By continuity, letting $t$ approach to infinity, $S \in \phi^{**}(N, e^S)$.

By step 3, $|S| = 2^k$ for $k \in \mathbb{N}$.

Step 5: For any coalition formation game $(N, \pi)$

$$\phi^{**}(N, \pi) = \arg \min_{M \in Q \cup \{N\}} |M|$$

where $Q(N, \pi) = \{S \in 2^N \text{ such that } S \in W_{(N,\pi)} \text{ and } |S| = 2^m \text{ for some } m \in \mathbb{N}\}$.

Proof.

By step 4, we know that $\phi^{**}(N, \pi) \subset Q(N, \pi)$.

Let $S$ be a coalition of the smallest size in $Q(N, \pi)$. Since the sharing rule equalizes the shares of the agents, then $S$ is also a coalition of the smallest size in $Q(N, \frac{\pi}{It} + e^S)$ for any $t \geq 1$.

By steps 2 and 3, there exists $t^* \in \mathbb{N}$ large enough such that $S \in \phi^{**}(N, \frac{\pi}{It^*} + e^S)$.

By SI, $S \in \phi^{**}(N, \pi + (It^*)e^S) = \phi^{**}(N, \pi + I(t^* - 1)e^S + \frac{I}{|S|}e^S)$.

Since $S$ is a coalition of the smallest size in $Q(N, \pi + I(t^* - 1)e^S)$, then $S$ will maximize the payoff of the agents among ESE-coalitions. Thus, by DER, $S \in \phi^{**}(N, \pi + I(t^* - 1)e^S)$.

Continuing similarly for rounds $t^* - 1, t^* - 2, \ldots, 0$, we have that $S \in \phi^{**}(N, \pi)$.

Finally, by DER, all chosen coalitions should give the same share of the resource to the agents. Hence, by step 1 they are of the same size. Therefore, $\phi^{**}(N, \pi)$ contain only coalitions of the minimal size in $Q(N, \pi)$.

**Proof of part** ***ii*****.**

Step 1. If agent $i$ is a dictator in the game $(N, \pi)$, that is $\pi_i > \pi(N \setminus i)$, then $\phi(N, \pi) = \{\{i\}\}$.

Proof. First notice that by IZ, $\phi(N, [\pi_i, 0_{N\setminus i}]) = \{\{i\}\}$, where $0_{N\setminus i}$ is the zero vector in $\mathbb{R}^{N\setminus i}$. Therefore, by continuity $i \in \phi(N, [\pi_i, \frac{1}{1+k^* \frac{I}{\pi_i}} \pi_{N\setminus i}])$ for $k^*$ large enough. By SI, $i \in \phi(N, [\pi_i + (k^*)I, \pi_{N\setminus i}])$. By DER, $i \in \phi(N, [\pi_i + (k^* - 1)I, \pi_{N\setminus i}])$ because coalition $\{i\}$ is the smallest ESE-coalition at the power profile $[\pi_i + (k^* - 1)I, \pi_{N\setminus i}]$. Continuing similarly $k^* - 2$ times, $i \in \phi(N, [\pi_i + (k^* - 2)I, \pi_{N\setminus i}])$, $i \in \phi(N, [\pi_i + (k^* - 3)I, \pi_{N\setminus i}])$, $\ldots$, $i \in \phi(N, [\pi_i, \pi_{N\setminus i}])$. Moreover, by DER, $\phi(N, \pi) = \{\{i\}\}$ because $i$'s payoff is greater by getting the prize alone instead of sharing it with another agent.

Step 2. Consider the game for three agents $(\{1, 2, 3\}, [\pi_1, \pi_2, \pi_3])$ such that $\pi_1 > \pi_2 > \pi_3$ and $\pi_1 < \pi_2 + \pi_3$. Then, $\phi(\{1, 2, 3\}, [\pi_1, \pi_2, \pi_3]) = \{\{1, 2, 3\}\}$.

Proof. We prove this by contradiction. First, notice that any winning coalition contains at least two agents. Suppose that coalition $S$ such that $|S| = 2$ is chosen, that is, $S \in \phi(\{1, 2, 3\}, [\pi_1, \pi_2, \pi_3])$. Then, by ESE, $S \in \phi(\{1, 2, 3\}, [(1 + k^* \frac{I}{\pi(S)})\pi_S, \pi_{-S}])$. For a large $k^*$, the game $(\{1, 2, 3\}, [(1 + k^* \frac{I}{\pi(S)})\pi_S, \pi_{-S}])$ has a dictator. Hence, by step 1, $S$ such that $|S| = 2$ cannot be chosen. This is a contradiction.

Finally, consider the sequence of coalition formation games $(\{1, 2, 3\}, [1 + \frac{1}{k}, 1, \frac{2}{k}])$ for $k = 1, 2, 3, \ldots$. By step 2, $\{1, 2, 3\} \in \phi(\{1, 2, 3\}, [1 + \frac{1}{k}, 1, \frac{2}{k}])$. Also, $\{1, 2, 3\}$ is winning in $(1, 1, 0)$. Hence, $\{1, 2, 3\} \in \phi(\{1, 2, 3\}, [1, 1, 0])$. This contradicts IZ.

***Proof of Proposition 4***

**Proof.** SI is satisfied since the min function is scale invariant.

To prove that the sequential transition correspondence $\bar{\phi}$ satisfies ESE, let $S^j \in \bar{\phi}(N, \pi)$ and suppose $S^j \neq N$. Because we restrict on feasible sequences of coalitions, we know that the only possible deviations from $S^j$ are only into feasible subsets. Since by definition of $\bar{\phi}$, there does not exist a coalition $S^m \subset S^j$ such that $S^m \in W_{(S^j, \pi_{S^j})}$ and since under proportional sharing, the relative power of every agent in $S^j$ is unchanged and the relative power of the agents in $N \setminus S^j$ goes to zero, then coalition $S^j$ will be chosen in further rounds. If $S^j = N$, with proportional sharing the relative powers are the same at every round and thus $S^j$ has to be chosen again in every round.

To show that DER is satisfied, suppose $S^j \in \bar{\phi}(N,\pi)$ and there exist another coalition $Z$ such that $Z \in W_{(N,\pi)}$ and $Z \in \bar{\phi}(N, \pi + I \cdot PR^N(Z,\pi))$ and $Z \notin \bar{\phi}(N,\pi)$. Since $Z \in \bar{\phi}(N, \pi + I \cdot PR^N(Z,\pi))$, then $Z = S^m$ for some $m$. By the definition of a feasible sequence of coalitions, $S^j \subset S^m$ or $S^m \subset S^j$. If $S^j \subset S^m$, then $PR(N, \pi^{\infty}_{S^m}) < PR(N, \pi^{\infty}_{S^j})\ \forall i \in S^j \cap S^m$. If $S^m \subset S^j$ and $S^m \notin \bar{\phi}(N,\pi)$, then there exists $S^h \subset S^m$ such that $S^h \in W_{(S^m, \pi_{S^m})}$. Hence, $S^m \notin \bar{\phi}(N, \pi + I \cdot PR^N(S^m,\pi))$ since $S^h$ will be a winning coalition after enough rounds are played.

■

# References


[1] D. Acemoglu, G. Egorov, and K. Sonin. Coalition Formation in Non-Democracies. *Review of Economic Studies*, 75(4):987–1009, 2008.

[2] D. Acemoglu, G. Egorov, and K. Sonin. Do Juntas Lead to Personal Rule? *American Economic Review: Papers and Proceedings*, 99(2):298–303, 2009.

[3] D. Acemoglu, M. Golosov, and A. Tsyvinski. Power Fluctuations and Political Economy. *Journal of Economic Theory*, 146(3):1009–1041, 2011.

[4] K.H. Baik. Winner-Help-Loser Group Formation in Rent-Seeking Contests. *Economics and Politics*, 6(2):147–162, 1994.

[5] R. Barro. The Control of Politicians: An Economic Model. *Public Choice*, 14(1):19–42, 1973.

[6] B. Bartling and F.A. von Siemens. Equal Sharing Rules in Partnerships. *Journal of Institutional and Theoretical Economics*, 166(2):299–320, 2010.

[7] F. Bloch and B. Dutta. Formation of Networks and Coalitions. *Mimeo University of Warwick*, 2010.

[8] A. Bogomolnaia and M. Jackson. The Stability of Hedonic Coalition Structures. *Games and Economic Behavior*, 38(2):201–230, 2002.

[9] B. Bueno de Mesquita, A. Smith, R. Siverson, and J. Morrow. *The Logic of Political Survival*. The MIT Press, Cambridge MA, 2003.

[10] M. Chwe. Farsighted Coalitional Stability. *Journal of Economic Theory*, 63(2):299–325, 1994.

[11] E. Diamantoudi and L. Xue. Farsighted Stability in Hedonic Games. *Social Choice and Welfare*, 21(1):39–61, 2003.

[12] L. Ehlers. Coalitional Strategy-proof House Allocation. *Journal of Economic Theory*, 105(2):298–317, 2002.

[13] J. Ferejohn. Incumbent Performance and Electoral Control. *Public Choice*, 50(1):5–25, 1986.

[14] E. Goode. Insane or Just Evil? A Psychiatrist Takes a New Look at Hitler. *The New York Times*, November. http://www.nytimes.com/1998/11/17/science/insane-or-just-evil-a-psychiatrist-takes-a-new-look-at-hitler.html. Accessed: 2015-08-14.

[15] E. Goode. The World; Stalin to Saddam: So Much for the Madman Theory. *The New York Times*, May. http://www.nytimes.com/2003/05/04/weekinreview/the-world-stalin-to-saddam-so-much-for-the-madman-theory.html. Accessed: 2015-08-14.

[16] K. Jandoc and R. Juarez. Self-Enforcing Coalitions under a General Class of Sharing Rules. *Mimeo University of Hawai'i*, 2015.

[17] J.S. Jordan. Pillage and Property. *Journal of Economic Theory*, 131(1):26–44, 2006.

[18] R. Juarez. Three Equilibria in Tournament Games with Externalities. *Mimeo University of Hawaii at Manoa*, 2012a.

[19] R. Juarez. The No-Threat Equilibrium in Tournament Games with Externalities. *Mimeo University of Hawaii at Manoa*, 2012b.

[20] R Juarez. Group Strategyproof Cost Sharing: The Role of Indifferences. *Games and Economic Behavior*, 82:218–239, 2013.

[21] H. Konishi and D. Ray. Coalition Formation as a Dynamic Process. *Journal of Economic Theory*, 110(1):1–41, 2003.

[22] S. Lee. Inter-Group Competition For A Pure Private Rent. *The Quarterly Review of Economics and Finance*, 33(3):251–266, 1993.

[23] S. Lee. Endogenous Sharing Rules in Collective-Group Rent Seeking. *Public Choice*, 85(1):31–44, 1995.

[24] M. McGuire and M. Olson. The Economics of Autocracy and Majority Rule: The Invisible Hand and the Use of Force. *Journal of Economic Literature*, 34(1):72–96, 1996.

[25] S. Nitzan. Collective Rent Dissipation. *Economic Journal*, 101(409):1522–1534, 1991.

[26] S. Nitzan. Modelling Rent Seeking Contests. *European Journal of Political Economy*, 10(1):41–60, 1994.

[27] M. Olson. *The Logic of Collective Action.* Harvard University Press, Cambridge, MA, 1965.

[28] M. Olson. Dictatorship, Democracy, and Development. *The American Political Science Review*, 87(3):567–576, 1993.

[29] S. Papai. Unique Stability in Simple Coalition Formation Games. *Games and Economic Behavior*, 48(2):337–354, 2004.

[30] G. Roland Persson, T. and G. Tabellini. Separation of Powers and Political Accountability. *The Quarterly Journal of Economics*, 112(4):1163–1202, 1997.

[31] G. Roland Persson, T. and G. Tabellini. Comparative Politics and Public Finance. *Journal of Political Economy*, 108(6):1121–1161, 2000.

[32] M. Piccione and R. Razin. Coalition Formation under Power Relations. *Theoretical Economics*, 4(1):1–15, 2009.

[33] D. Ray. *A Game-Theoretic Perspective on Coalition Formation.* Oxford University Press, New York, 2007.

[34] D. Ray and R. Vohra. *Coalition Formation.* North-Holland, 2013.

[35] D. Ray and R. Vohra. The Farsighted Stable Set. *Econometrica*, 83(3):977–1011, 2015.

[36] P. Sekeris. Endogenous Elites: Power Structure and Patron-Client Relationships. *Economics of Governance*, 12(3):237–258, 2011.

[37] W. Thomson. Axiomatic and Game-theoretic Analysis of Bankruptcy and Taxation Problems: a Survey. *Mathematical Social Science*, 45(3):249–297, 2003.

[38] W. Thomson. On the Existence of Consistent Rules to Adjudicate Conflicting Claims: a Constructive Geometric Approach. *Review of Economic Design*, 11(3):225–251, 2007.

[39] G. Tullock. *Autocracy.* Kluwer Academic Publishers, Dorderecht, 1987.

[40] L. Xue. Coalitional Stability Under Perfect Foresight. *Economic Theory*, 11(3):603–627, 1998.