Coalition Formation Under Power Relations*

Michele Piccione and Ronny Razin
Department of Economics
London School of Economics

This version: March 2008

Abstract

We analyze the structure of a society driven by power relations. Our model has an exogenous individual power relation between agents and a corresponding coalitional power relation between coalitions of agents. Agents determine the social order by forming coalitions. The power relations determine the ranking of agents in society for any social order. We study a cooperative game in partition form and introduce a solution concept, the *stable social order*, which exists and includes the core. We investigate a refinement, the *strongly stable social order*, which incorporates a notion of robustness to variable power relations. We provide a complete characterisation of strongly stable social orders.

*We wish to thank John Moore and Ariel Rubinstein. We are grateful to William Lauderdale and Iddo Tuvnel for their support.

1 Introduction

Power relations are a fundamental component of human interaction. In social environments, two types of power shape a significant number of human relations: individual power and group power. Individual power manifests itself in one-to-one relations and generally originates from material or psychological strength. Group power manifests itself in interactions between coalitions of individuals or in one-to-one interactions between individuals belonging to different coalitions. The objective of this paper is to study theoretically the joint influence of individual and group power in the determination of social arrangements. Although the term "individual" usually refers to "one person", in this paper "individuals" can be entities such as families, factions or other groupings, the unity of which is solid and based on exogenous, non-strategic factors such as blood, loyalty, or friendship. Henceforth, such individuals or families will be referred to as "agents".

We are interested in analyzing the structure of a society driven by power relations. Our model has the following basic ingredients. The primitives are an exogenous individual power relation over agents and a coalitional power relation consistent with the individual power relation. Agents determine the social order by forming coalitions. The power relations determine the ranking of agents in society for any social order.

Coalitions, in our model, are held together only by strategic considerations. We assume that the objective of each agent is to maximize his/her position in the societal ranking. We study a cooperative game in partition form and introduce a solution concept, the *stable social order*. We show that a stable social order exists and includes the core. We investigate a refinement, the *strongly stable social order*, which incorporates a stringent notion of robustness to variable power relations. We provide a complete characterisation of strongly stable social orders.

Our framework is too abstract to fit specific historical examples. However, several implications of our results are broadly consistent with stylized historical and political anecdotes. In particular, in a strongly stable social order:

(1) Powerful coalitions are large and each coalition is immune from the threat of a unified challenge coming from all less powerful coalitions.

(2) Leaders are critical. The elimination of society's most powerful member causes a regime switch: almost all the members of the coalition in power divide into smaller coalitions and significantly drop in status.

As we shall see, the robustness criterion implicit in strongly stable social orders is rather demanding. Hence, we conclude this paper by focusing on social orders that are stable (not necessarily strongly stable) for special power relations.

1.1 A Simple Example

Suppose that the individual power of agent i is represented by a number $q(i)$: agent i is more powerful than agent j if and only if $q(i) > q(j)$. The power relation between any two disjoint coalitions of agents is determined additively, i.e., coalition A is more powerful than coalition B whenever $\sum\limits_{i \in A} q(i) > \sum\limits_{i \in B} q(i)$.

Suppose that all numbers $q(i)$ are decreasing in i, that is, agent 1 is the most powerful, agent 2 is the second most powerful, and so on. Also suppose that all numbers $q(i)$ are *approximately the same,* that is, a coalition of m agents is more powerful than any coalition with less than m agents, and that no two coalitions have the same power. Consider the following partition (social order) $\hat{\Sigma}$ for the set of agents $I = 1, 2, 3, 4, 5, 6, 7$,

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

The social order modifies the exogenous individual power ranking in the following way. First, the agents in a stronger coalition are more powerful than the agents in a weaker coalition. Second, within a coalition, agents are ranked according to their individual power. Thus, agent 1 is the most powerful, agent 3 is the second most powerful, agent 2 is the fifth most powerful and so on.

Suppose that agents care only about their ranking in the social order. We say that a social order is in the *core* if there does not exist a subset of agents who can strictly improve their position in the ranking by forming a new coalition. The above social order is not in the core. If agents 3, 5, and 7 form a new coalition C' dropping agent 1, they strictly improve their position in the resulting social order

$$\hat{\Sigma}' = \{\{3, 5, 7\}, \{2, 6\}, \{1\}, \{4\}\}.$$

Piccione and Rubinstein (2004) show that the core is empty when $N > 6$. In Section 5 we show that other standard solution concepts also fail to guarantee existence. Our aim is to provide a solution concept for which existence is not problematic and which can offer interesting insights into coalition formation in the presence of power relations. We follow the traditional route of restricting the set of profitable deviations. Our *stable social order* incorporates two features:

(i) A recursive definition of *"durable"* deviations and counter-deviations.
(ii) A sequential notion of counter-deviations: members of a deviating coalition do not participate in any immediately subsequent counter-deviation.

The social order $\hat{\Sigma}$ is stable. In particular, (all) members of the coalition C' are made worse off by the counter-deviation $C'' = \{1, 2, 4, 6\}$ which is "durable" in that agents $1, 2, 4, 6$ cannot subsequently be made worse off by any coalition of agents not in C''.

As we shall see, the social order $\hat{\Sigma}$ is also *strongly stable*, that is, it is stable for any selection of the numbers $q(i)$ that are decreasing in i; irrespective of the cardinal properties of $q()$, the agents in C' are made worse off by the counter-deviation C'' and the agents in C'' cannot be made worse off.

1.2 Related literature

This paper is obviously part of the vast literature on cooperative games, solution concepts, and coalition formation. We refer the reader to textbooks such as Myerson (1991), Osborne and Rubinstein (1994), or Ray (2007) for a detailed and insightful overview. Games in partition forms were studied in Lucas and Thrall (1963), Myerson (1977), Ray and Vohra (1999). Our solution concept is related to the notion of the "Bargaining Set" of Aumann and Maschler (1964) and to notions of "farsightedness" developed in Chwe (1994), Ray and Vohra (1997), Greenberg (2000), Diamantoudi and Xue (2007). Formal models of power relations were analysed in Jordan (2006a,b), Piccione and Rubinstein (2007), and Acemoglu, Egorov, and Sonin (2007). Jordan (2006a) considers a model in which power is endogenous and is affected by the wealth that is appropriated from other agents through the exercise of power. Jordan (2006b) incorporates dynamic factors such as histories into the notion of stability, thus introducing a notion of "legitimacy" into the appropriation process. Piccione and Rubinstein (2007) study a model in which the allocation of resources is driven by exogenous power. In this paper, we report a result from Piccione and Rubinstein (2004) which was omitted in Piccione and Rubinstein (2007). Acemoglu, Egorov, and Sonin (2007) also assume that power is exogenous and study the formation of coalitions under an allocation rule for which the winning coalition takes all.

2 The model

The set of agents is $I = \{1, ...N\}$. Although the term "agent" is commonly associated with "one person", in our model an agent can be a clan, a family, or any group of people held together by non-strategic factors. The agents are ordered by an exogenous, individual power relation P. The binary relation P is irreflexive, asymmetric, complete, and transitive, and without loss of generality, is such that $1P2$,$2P3$,...,$(N-1)PN$. The statement "i is exogenously more powerful than j" is denoted by iPj.

Power is endogenously redistributed through the formation of coalitions. We define a "coalitional" power relation over sets of agents as a binary relation Π between subsets (coalitions) of agents $A, B \subset I$ such that $A \cap B = \varnothing$. The relation Π is asymmetric, acyclic[1], and such that either $A\Pi B$ or $B\Pi A$ whenever $A \cap B = \varnothing$. The statement $A\Pi B$ is interpreted as "coalition A is more powerful than coalition B". We assume that $A\Pi\varnothing$ whenever $A \subset I$ is non-empty. Note that two disjoint coalitions cannot be equally powerful.

The structure of coalitions and the power relations jointly determine the ranking of agents with respect to power. We assume that each agent strictly prefers a higher position in the power ranking to a lower position.[2]

We define a *social order* as a partition of agents. Formally, a social order is a partition $\Sigma = \{C_1, ..., C_K\}$ of the set I. We adopt the convention that $C_i\Pi C_j$ if and only if $i < j$.

We will restrict our analysis to coalitional power relations that are 'consistent' with the individual power relation P. The coalitional power relation

[1] The relation Π is acyclic if, given any collection Θ of subsets of agents, there exists $A \in \Theta$ such that $B\Pi A$ for no $B \in \Theta$.

[2] This formulation is quite general. In a more concrete example, the endowment to be distributed is a set of houses $H = \{1, ..., N\}$. Each agent i can consume only one house, all share the same ordering over houses, and strictly prefer having any house to having no house. The houses are chosen sequentially by the agents in accordance with the ranking.

Π is *aggregation consistent* if, for any subsets of agents A_1, A_2, A_3, A_4 such that $A_i \cap A_j = \varnothing$, $i \neq j$, $A_1 \Pi A_3$ and $A_2 \Pi A_4$ implies that

$$(A_1 \cup A_2) \, \Pi \, (A_3 \cup A_4)$$

We say that a coalitional power relation Π is $P-$consistent if it is aggregation consistent and the restriction of Π on singleton coalitions is P. Consider two coalitions of agents $A, B \subset I$ such that $A \cap B = \varnothing$. Coalition A *dominates* coalition B if there exists a one-to-one mapping $\sigma : B \longrightarrow A$ such that $iP\sigma^{-1}(i)$ for any $i \in \sigma(B)$. The next Lemmas will be useful later.

Lemma 1 *Suppose* Π *is aggregation consistent. Then, if* $A \Pi B$ *and* $C \subset B$, *then* $A \Pi C$.

Proof: Suppose not. A contradiction is obtained defining $A_1 = C$, $A_2 = B/C$, $A_3 = A$, $A_4 = \varnothing$. QED

Lemma 2 *Suppose* Π *is* $P-$*Consistent. Then* $A \Pi B$ *whenever* A *dominates* B.

Proof: It follows from a simple application of Lemma 1. QED

Consider a coalitional power relation Π. The relation Π is *monotonic* if for any two subsets of agents A_1, A_2 such that $A_1 \cap A_2 = \varnothing$ and $A_1 \Pi A_2$, implies that $((A_1 \cup \{i\})/\{j\}) \, \Pi \, ((A_2 \cup \{j\})/\{i\})$ whenever $i \in A_2$ and iPj.

Lemma 3 *If* Π *is aggregation consistent, then* Π *is monotonic.*

Proof: Consider two subsets of agents A_1, A_2 such that $A_1 \cap A_2 = \varnothing$ and $A_1 \Pi A_2$. Take $i \in A_2$ and any j such that iPj. First suppose that $(A_2/\{i\}) \, \Pi \, (A_1/\{j\})$. Then, by aggregation consistency, $A_2 \Pi \left((A_1/\{j\}) \cup \{j\}\right)$. Since $A_1 \subset (A_1/\{j\}) \cup \{j\}$, a contradiction is obtained by Lemma 1. Hence, $(A_1/\{j\}) \, \Pi \, (A_2/\{i\})$. The claim follows by aggregation consistency. QED

Throughout this paper, we will maintain the assumption of $P-$consistency.

Finally, we define the power relation that is induced by a social order. Given a coalitional power relation Π that is $P-$consistent and a social order $\Sigma = \{C_1, ..., C_K\}$, we define the *induced* power relation $Q^{\Pi}(\Sigma)$ on I as follows:

(*) $iQ^{\Pi}(\Sigma)j$ if and only if

either $i, j \in C_k$ for some k and iPj

or $i \in C_k$, $j \in C_{k'}$ and $C_k \Pi C_{k'}$.

3 Stability

Let $V_i^{\Pi}(\Sigma)$ denote agent i's position in the ranking induced by $Q^{\Pi}(\Sigma)$, that is, $V_i^{\Pi}(\Sigma) = 1$ indicates that agent i is the most powerful, $V_i^{\Pi}(\Sigma) = 5$ indicates that agent i is the fifth most powerful, and so on. We introduce a cooperative solution concept for social orders that we call *stable social order*.

For any subset of agents C and a social order Σ, with some abuse of the conventional notation let $\Sigma \top C$ be the partition $\{C_1/C, C_2/C, ..., C_K/C, C\}$. We define the following stability notion.

Let Ξ be the set of social orders and $\mathcal{I}$ be the set of all possible subsets of I. Define the correspondence $\mathcal{S}^{\Pi} : \Xi \Longrightarrow \mathcal{I}$ such that $C \in \mathcal{S}^{\Pi}(\Sigma)$ if and only if

(a) $V_i^{\Pi}(\Sigma \top C) < V_i^{\Pi}(\Sigma)$ for any $i \in C$; and

(b) there does not exist $C' \in \mathcal{S}^{\Pi}(\Sigma \top C)$ such that

  (i) $C \cap C' = \varnothing$;

  (ii) $V_i^{\Pi}((\Sigma \top C) \top C')) > V_i^{\Pi}(\Sigma)$ for some $i \in C$;

A deviation C from a social order Σ is *durable* if $C \in \mathcal{S}^{\Pi}(\Sigma)$. Two criteria need to be satisfied by a durable deviation. First, all members in the deviating coalition are better off. Second, there does not exist a durable counter-deviation that makes some member in the original deviating coalition worse off than in the social order prior to the deviation. It should be noted that members in the deviating coalition are excluded from the counter-deviating coalition.

The following proposition shows that the mapping $\mathcal{S}^{\Pi} : \Xi \Longrightarrow \mathcal{I}$ exists and is unique notwithstanding its self-referential nature.

Proposition 4 *There exists a unique correspondence $\mathcal{S}^{\Pi} : \Xi \Longrightarrow \mathcal{I}$ which satisfies (a) and (b).*

Proof: Given a social order Σ and a coalition C such that $V_i^{\Pi}(\Sigma \top C) < V_i^{\Pi}(\Sigma)$ for any $i \in C$, let

$$\tilde{\mathcal{S}}(\Sigma, C) = \{C' \subset I : \begin{array}{l} \text{(i) } C \cap C' = \varnothing \\ \text{(ii) } V_i^{\Pi}((\Sigma \top C) \top C')) < V_i^{\Pi}(\Sigma \top C) \text{ for each } i \in C' \\ \text{(iii) } V_i^{\Pi}((\Sigma \top C) \top C')) > V_i^{\Pi}(\Sigma) \text{ for some } i \in C\} \end{array}$$

Consider all finite sequences $\{B^t\}_{t=0}^{\tau}$ of subsets of agents such that

- $B^0 = C$;
- $\Sigma^0 = \Sigma$, $\Sigma^{t+1} = \Sigma^t \top B^t$;
- $B^t \in \tilde{\mathcal{S}}(\Sigma^{t-1}, B^{t-1})$, $B^t \cap B^{t-1} = \varnothing$ for $t > 0$;
- Either $\tilde{\mathcal{S}}(\Sigma^{t-1}, B^{t-1}) \neq \varnothing$ for any $t \leq \tau$ and $\tilde{\mathcal{S}}(\Sigma^{\tau}, B^{\tau}) = \varnothing$ or $\tilde{\mathcal{S}}(\Sigma^{t-1}, B^{t-1}) \neq \varnothing$ for any t and $\tau = \infty$.

Note that, by (ii) and (iii) in the definition of $\tilde{\mathcal{S}}$, each member of B^t is better off in $\Sigma^{\tau+1}$ than in Σ^{τ} and that at least one member of B^t is worse off in $\Sigma^{\tau+2}$ than in Σ^{τ}. Hence, $B^t \Pi B^{t-1}$ for every $t > 0$. Thus, by acyclicity, there exists a finite bound for τ that is common to all sequences $\{B^t\}_{t=0}^{\tau}$. Since $\tilde{\mathcal{S}}(\Sigma^{\tau}, B^{\tau}) = \varnothing$, if $B \in \mathcal{S}^{\Pi}(\Sigma^{\tau} \top B^{\tau})$ then $B \cap B^{\tau} \neq \varnothing$. Hence, $B^{\tau} \in \mathcal{S}^{\Pi}(\Sigma^{\tau})$ and $B^{\tau-1} \notin \mathcal{S}^{\Pi}(\Sigma^{\tau-1})$. Consider now a directed graph in which each B^t is a node and a directed edge links B^t to B^{t+1} if and only if B^t immediately precedes B^{t+1} in the same sequence. If none of the immediate successors of B^t is in $\mathcal{S}^{\Pi}(\Sigma^t \top B^t)$, then $B^t \in \mathcal{S}^{\Pi}(\Sigma^t)$. If at least one immediate successor of

B^t is in $\mathcal{S}^{\Pi}\left(\Sigma^t \top B^t\right)$, then $B^t \notin \mathcal{S}^{\Pi}\left(\Sigma^t\right)$. Proceeding by backward induction in this fashion, we determine uniquely whether $C \in \mathcal{S}^{\Pi}(\Sigma)$. QED

The example in Section 1.1 clarifies the intuition behind this result. The coalition C'' is a durable deviation from $\hat{\Sigma}'$ since the agents in C'' cannot be made worse off by any coalition of agents that are not in C''. Hence, working backwards, one obtains that the coalition C' is not a durable deviation from $\hat{\Sigma}$.

We are now ready to define the stability of a social order:

Definition: A social order $\Sigma = \{C_1, .., C_K\}$ is *stable* with respect to a coalitional power relation Π if $\mathcal{S}^{\Pi}(\Sigma) = \varnothing$.

A social order is stable if it cannot be upset by a durable deviation, or equivalently, any deviation from the social order has a durable counter-deviation. It should be noted that durability is defined for off-equilibrium behavior. Notably, a social order resulting from a durable deviation is not necessarily stable. As customary in non cooperative game theory and in many solution concepts in cooperative game theory, we treat deviations and equilibria asymmetrically.

4 The main result

In this section, we introduce and prove our main result.

Theorem 5 *There exists a social order that is stable for any $P-$consistent Π.*

The proof of Theorem 5 is constructive. We first define a social order Σ^*. We then prove that Σ^* is stable for any $P-$consistent Π.

4.1 The social order Σ^*

The social order Σ^* is constructed with a simple procedure. First, select the odd-indexed agents to form the strongest coalition. Re-index the remaining agents so that the most powerful agent is indexed as agent $1'$, the second most powerful is indexed as agent $2'$ and so on. Select the odd indexed agents from this set to form the second strongest coalition. Repeat this procedure until no agents are left.

Formally, Let $\lfloor d \rfloor$ denote the largest integer smaller than or equal to d and $\lceil d \rceil$ the smallest integer larger than or equal to d. We construct the social order $\Sigma^* = \{C_1^*, ..., C_K^*\}$ recursively as follows.

1. First, let $C_1^* = \{1, 3, ..., 2\left\lfloor \frac{N+1}{2} \right\rfloor - 1\}$.
2. Now take $I/C_1^* = \{2, 4, ..., 2\left\lceil \frac{N-1}{2} \right\rceil\}$ and temporarily re-lable the agents in I/C_1^* in decreasing power as $1', 2', .., , m_1'$, where m_1' is the number of agents in I/C_1^*.
3. Define $C_2^* = \{1', 3', ..., (2\left\lfloor \frac{m_1+1}{2} \right\rfloor - 1)'\}$ with the agents re-assuming their original index.
4. Suppose we have defined C_j^* for $j = 1, 2, ..., k$. Take $I/\{\cup_{j=1}^k C_j^*\}$ and re-lable its agents by decreasing power as $1', 2', .., m_j'$, where m_j' is the number of agents in $I/\{\cup_{j=1}^k C_j^*\}$.
5. Define $C_{k+1}^* = \{1', 3', ..., (2\left\lfloor \frac{m_j+1}{2} \right\rfloor - 1)'\}$ with the agents re-assuming their original index.
6. The procedure ends when $I/\{\cup_{j=1}^k C_j^*\} = \phi$.

4.2 Stability

We use the claims below to prove Theorem 5.

Claim 0: For any j, $C_j^* \Pi (\cup_{i=j+1}^K C_i^*)$.
Proof: By construction. ■

Claim 1: Fix some partition Σ such that $C_1^* \in \Sigma$ and $2 \in C_2$. For any C such that $C_1^* \cap C \neq \varnothing$ and $V_i^{\Pi}(\Sigma \top C) < V_i^{\Pi}(\Sigma)$ for each $i \in C$, there exists a C', $C' \cap C = \varnothing$, such that

(i) $V_i^{\Pi}((\Sigma^* \top C) \top C')) < V_i^{\Pi}(\Sigma^* \top C))$ for each $i \in C'$;

(ii) $V_i^{\Pi}((\Sigma^* \top C) \top C')) > V_i^{\Pi}(\Sigma^*)$ for some $i \in C$;

(iii) $C' \Pi((I/C)/C')$.

Proof: Denote $C_1^* = \{y_1, ..., y_L\}$, $C = (x_1, ..., x_M)$. Construct $C' = \{z_1, ..., z_L\}$ by first letting $z_1 = y_1 = 1$. Suppose we have defined z_i for all $i \leqslant j$ for some $j \geqslant 1$. Define z_{j+1} as the smallest i such that: (i) $i \notin C$; (ii) $i \neq z_1, ..., z_j$; (iii) $V_i^{\Pi}(\Sigma \top C) > j+1$; (iv) $i \leq y_{j+1}$. We now show that the above algorithm is well defined. We consider several cases:

Case 1: In $\Sigma \top C$, C is ranked first and C_1^*/C is ranked second.

First note that either $z_2 = 2$, or $2 \in C$ and therefore $3 \notin C$ implying $z_2 = 3$. Therefore, $z_2 \leqslant y_2$. Now consider z_j, $j > 2$, given that z_1,...,z_{j-1} which have been selected using the above algorithm. Let G_j be the set of agents smaller than or equal to $y_j = 2j - 1$. By hypothesis, $j - 1$ agents in G_j have already been allocated to C'. We now show that the set $H_j = \{r \in G_j/\{z_1, ..., z_{j-1}\} : r \notin C \text{ and } V_r^{\Pi}(\Sigma \top C) > j\}$ is not empty. Since $\#G_j/\{z_1, ..., z_{j-1}\} = j$, it is impossible that all agents in $G_j/\{z_1, ..., z_{j-1}\}$ are in C. If so, agent $2j - 1$ would also be in C but ranked at or worse then the j^{th} position in $\Sigma \top C$, contradicting the definition of C. Therefore, there must exist at least an agent $i \in G_j/\{z_1, ..., z_{j-1}\}$ such that $i \notin C$. Consider then the agent r^* in $G_j/\{z_1, ..., z_{j-1}\}$ that is ranked worst in $\Sigma \top C$. Agent r^* is not in C as otherwise $G_j/\{z_1, ..., z_{j-1}\} \subset C$. Since C_1^*/C is ranked second, agent r^* must be ranked worse than agent 1 in $\Sigma \top C$. Since agent 1 is not in $G_j/\{z_1, ..., z_{j-1}\}$ and $\#G_j/\{z_1, ..., z_{j-1}\} = j$, $V_{r^*}^{\Pi}(\Sigma \top C) > j$. Hence, H_j is not empty. Define $z_j = \min H_j$. ■

Case 2 In $\Sigma \top C$, C is ranked first and C_1^*/C is ranked worse than second.

Again, either $z_2 = 2$, or $2 \in C$ and therefore $3 \notin C$ implying $z_2 = 3$. Consider z_j, $j > 2$. As in Case 1, $\#G_j/\{z_1, ..., z_{j-1}\} = j$. Denote the agents in $G_j/\{z_1, ..., z_{j-1}\}$ by $\gamma_1 < ... < \gamma_j$. We need to show that in $\Sigma \top C$ one γ_k is ranked strictly worse than j^{th} position. If agent $2j-1$ is not in C, then the claim is obvious as C_1^*/C is ranked worse than second. Hence, we can suppose that agent $2j-1$ is in C. Note that it is impossible that all agents $\gamma_1, ..., \gamma_j$ are in C. If so, agent $\gamma_j = 2j-1$ is ranked at most j^{th} in C, contradicting the definition of C. Also, if agent $2j-1$ is not the lowest ranked agent in C, agents not in $\gamma_1, ..., \gamma_j$ are also in C. Since not all agents $\gamma_1, ..., \gamma_j$ are in C, one γ_k must be ranked strictly worse than the j^{th} position. We can then suppose that $2j-1$ is the worst ranked agent in C.

If C does not contain an even agent, then agent 2 is the most powerful agent in the second most powerful faction in $\Sigma \top C$, which we denote by $C_2^{\Sigma \top C}$. In this case, $z_2 = 2$ and agent 2 is not in $\gamma_1, ..., \gamma_j$. Since some agents in $\gamma_1, ..., \gamma_j$ are not in C and agent 2 is the most powerful agent in $C_2^{\Sigma \top C}$, one γ_k must be ranked strictly worse than the j^{th} position. Hence, if C does not contain an even agent, the claim is proven. Suppose then that C does contain an even agent. Note further that for this even agent γ, $\gamma < 2j-1$, as otherwise agent $2j-1$ is not the least powerful in C.

To summarise, in order to conclude the proof of Case 2, we suppose that

(i) not all $\gamma_1, ..., \gamma_j$ are in C;

(ii) $2j-1$ is the worst ranked agent in C;

(iii) C contains an even agent.

Also, if $C_2^{\Sigma \top C}$ contains an agent z_k, $k < j$, (i) implies that at least one γ_l must be ranked strictly worse than the j^{th} position in $\Sigma \top C$. If not, agent z_k is ranked strictly worse than the j^{th} position in $\Sigma \top C$, the agent γ_l that is ranked in the j^{th} in $\Sigma \top C$ is in $C_2^{\Sigma \top C}$ and, $\gamma_l < z_k$. As $j > k$, the algorithm should not have selected z_k. Hence, we also suppose that

(iv) $C_2^{\Sigma \top C}$ does not contain any agents in z_1,...,z_{j-1}.

Now let Θ be the set $C_1^* \cap C$. Given any $(2k-1) \in C$, $k < j$, if q odd agents smaller than or equal to $(2k-1)$ are in C, at least q even agents that are smaller than $(2k-1)$ are in $z_1, ..., z_{j-1}$. By (ii), at least $\#\Theta - 1$ even agents must be in $z_1, ..., z_{j-1}$. Call Θ' the set composed of these even agents and one even agent i' from C by (iii). We can now construct a one to one mapping $g : \Theta \to \Theta'$ such that $g(z) < z$. First, let $g(2j-1) = i' < 2j-1$ by (ii) and (iii). Let $\Theta = \{\theta_1, ..., \theta_m\}$, $\theta_i < \theta_{i+1}$. It is easy to see that there must be an even agent $i \in \Theta'/\{i'\}$ such that $i < \theta_1$. Define $g(\theta_1)$ as the lowest such number. Suppose that for $k-1$ agents θ_i, $i<k$, $k>1$, we have constructed $g(z)$. Since k odd agents smaller than or equal to θ_k are in C, there must exist k even agents in $z_1, ..., z_k$ that are smaller than θ_k. Hence, there must be an even agent $\theta'_k > g(\theta_{k-1})$, $\theta'_k \in \Theta'/\{i'\}$ and $\theta'_k < \theta_k$. Let $g(\theta_k)$ be the strongest of such agents.

Hence, Θ' dominates Θ. By (iv), $\Theta' \cap C_2^{\Sigma \top C} = \varnothing$. Since $C_1^* \Pi (N/C_1^*)$, then $(C_1^*/\Theta)\, \Pi((N/C_1^*)/\Theta')$ by aggregation consistency. Since $C_2^{\Sigma \top C}$ is contained in $(N/C_1^*)/\Theta'$, Lemma 1 implies $(C_1^*/\Theta)\, \Pi C_2^{\Sigma \top C}$, a contradiction. ■

Case 3 C is not ranked first in $\Sigma \top C$.

Let $C_1^{\Sigma \top C}$ be the most powerful coalition in $\Sigma \top C$. Since C_1^*/C is ranked worse than $C_1^{\Sigma \top C}$, $\#C_1^{\Sigma \top C} \geqslant 2$. Hence, agent 3 cannot be in C. Since agent 3 is in C_1^*/C, $\#C_1^{\Sigma \top C} \geqslant 3$. Continuing in this fashion, we establish that $C \bigcap C_1^* = \phi$, a contradiction .■

By construction, $V_i^{\Pi}((\Sigma^* \top C) \top C')) < V_i^{\Pi}(\Sigma^* \top C))$ for all $i \in C'$. To see that $V_i^{\Pi}((\Sigma^* \top C) \top C')) > V_i^{\Pi}(\Sigma^*)$ for some $i \in C$, take any $\hat{x} \in C \cap C_1^*$. Agent $\hat{x}$'s position is weakly better than the L^{th} position in Σ^* and strictly worse than the L^{th} position in $(\Sigma^* \top C) \top C')$.

Finally, we show that $C'\Pi((I/C)/C')$. Indeed, our construction ensures that $C'\Pi(I/C')$. By definition, $C_1^*\Pi(I/C_1^*)$. Since C' is derived from C_1^* exchanging less powerful agents in C_1^* for more powerful agents in I/C_1^*, monotonicity implies that $C'\Pi(I/C')$. QED

Proof of Theorem 5: Fix some partition Σ such that $C_1^*, C_2^* \in \Sigma$ and $\min C_3^* \in C_3$. For any C such that $C_1^* \cap C = \varnothing$, $C_2^* \cap C \neq \varnothing$ and $V_i^{\Pi}(\Sigma \top C) < V_i^{\Pi}(\Sigma)$ for each $i \in C$, construct a counter-deviation C'' which is constructed analogously to C' in Claim 1 (ignoring the agents in C_1^*). Namely, denoting $C_2^* = \{y_1, ..., y_{L'}\}$, $C = (x_1, ..., x_{M'})$, construct $C'' = \{z_1, ..., z_{L'}\}$ by first letting $z_1 = y_1 = 2$. Having defined z_i for all $i \leqslant j$ for some $j \geqslant 1$, define z_{j+1} as the smallest i such that: (i) $i \notin C$; (ii) $i \neq z_1, ..., z_j$; (iii) $V_i^{\Pi}(\Sigma \top C) > j + 1 + \#C_1^*$; (iv) $i \leq y_{j+1}$. The deviation C'' is durable; any counter-deviation to C'' in $(\Sigma \top C) \top C''$ cannot be durable by Claim 1.

The completion of the proof is obtained by an inductive repetition of these arguments. QED

5 Alternative notions of stability

In this section, we focus on other stability criteria such as the core and coalition proofness. We say that the relation Π is *homogeneous* if a coalition of m agents is more powerful than any coalition with strictly less than m agents.

A social order $\Sigma = \{C_1, ..., C_K\}$ is in the *core* if there does not exist a subset of agents C such that $V_i^{\Pi}(\Sigma \top C) < V_i^{\Pi}(\Sigma)$ for each $i \in C$.

We are interested in conditions under which the "core" is empty. The next proposition is from Piccione and Rubinstein (2004).

Proposition 6 *If Π is homogeneous, the core is empty when $N \geq 7$.*

Proof: Let $\#C_j$ be the number of agents in C_j.

Fact 1: The least powerful coalition C_K in Σ has either 1 or 2 agents.

Proof: If not, all agents in C_K except for the most powerful can form a coalition which strictly improves their ranking.

Fact 2: $\#C_{j+1} \leq \#C_j \leq \#C_{j+1} + 1$ for $j = 0, ..., K - 1$.

Proof: The left hand side follows by definition and by homogeneity. For the right hand side, if $\#C_j > \#C_{j+1} + 1$, all agents in C_j except for the most powerful can form a coalition which strictly improves their ranking.

Fact 3: $\#C_2 \geq 2$ and $\#C_1 \geq 3$.

Proof: If $\#C_2 = 1$ then, by Facts 1 and 2, $K > 5$ and $\#C_j = 1$ for $j = 2, 3, 4, .., K$. Thus, merging C_K and C_{K-1} improves the ranking of all the members of the new coalition. Hence, $\#C_2 \geq 2$. If $\#C_1 = 2$, then merging C_2 with one element of C_K improves the ranking of all the members of the new coalition.

Since $N \geq 7$, there are at least two agents who do not belong to either C_1 or C_2. If two such agents form a coalition with the agents in C_2, the ranking of all the members of the new coalition improves. QED

Note that, for $N = 6$, the social order $\{\{1, 5, 6\}, \{3, 4\}, \{2\}\}$ can be in the core when Π is homogeneous.

The set of coalition-proof social orders is sometimes empty as well. We adopt a definition that is stronger than necessary. A social order $\Sigma =$

$\{C_1, ..., C_K\}$ fails *Coalition-Proofness* if there exists a subset of agents C such that $V_i^{\Pi}(\Sigma \top C) < V_i^{\Pi}(\Sigma)$ for each $i \in C$ and for any subset of agents $C' \subset C$ there exists $i \in C'$ such that $V_i^{\Pi}((\Sigma \top C) \top C') \geq V_i^{\Pi}(\Sigma \top C)$.

We say that the function Π satisfies property MW if and only if for any subset of agents $C = \{i_1, i_2 ..., i_k\} \subset N$ such that k is even and $i_j P i_{j+1}$, $j = 1, .., k-1$, $\{i_2, i_3, ..., i_{\frac{k}{2}+1}\} \Pi \{i_1, i_{\frac{k}{2}+2}, ..., i_k\}$.

Remark: Suppose that a number $q(i)$ is associated with each agent i and iPj if and only if $q(i) > q(j)$. Now assume that $A \Pi B$ whenever $\sum_{i \in A} q(i) > \sum_{i \in B} q(i)$. The following function, $q\,()$, for $N = 7$ (generalisable to any N) satisfies MW: Let $q(1) = 1.111111$, $q(2) = 1.11111$, $q(3) = 1.1111$, and so on. The function q is such that, for any two equal-sized subsets of agents, the one with the least powerful agent is less powerful.

Proposition 7 *Suppose that $N \geq 7$. If Π is homogeneous and satisfies MW, any social order $\Sigma = \{C_1, ..., C_K\}$ fail Coalition-Proofness.*

Proof: Let $\#C_j$ be the number of agents in C_j.

Fact 1: The least powerful coalition C_K in Σ has either 1 or 2 agents.

Proof: Suppose not. If $\#C_K$ is odd, the $\frac{\#C_K - 1}{2} + 1$ weakest agents in C_K can form a coalition which strictly improves their ranking and have no incentive to split any further. If $\#C_K$ is even then the 2^{nd} to the $\left(\frac{\#C_K}{2} + 1\right)^{th}$ weakest agents can form a coalition which strictly improves their ranking by property MW. In such coalition, they have no incentive to split any further. $\blacksquare$

Fact 2: $\#C_{j+1} \leq \#C_j \leq \#C_{j+1} + 1$ for $j = 0, ..., K-1$.

Proof: The left hand side follows by definition and the fact that all agents have similar strength. For the right hand side, if $\#C_j > \#C_{j+1} + 1$, then:

case 1 ($\#C_{j+1} + 1 < \#C_j < 2\#C_{j+1}$): form a new coalition C that is a subset of C_j removing the strongest agents from C_j until removing an additional agent makes the agents left in C_j less powerful than C_{j+1}. There are no incentives for the agents in this coalition to split further. Since $\#C_j < 2\#C_{j+1}$, $C \Pi (C_j - C)$

case 2: ($\#C_j = 2\#C_{j+1}$): as in Fact 1 the 2^{nd} to the $\left(\frac{\#C_j}{2}+1\right)^{th}$ weakest agents can form a coalition which strictly improves their ranking by property MW. In such coalition, they have no incentive to split any further.

case 3: $\#C_j > 2\#C_{j+1}$ we use the same argument as in Fact 1 since the new coalition will obviously be ranked better than C_{j+1}. $\blacksquare$

Fact 3: $\#C_2 \geq 2$ and $\#C_1 \geq 3$.

Proof: If $\#C_2 = 1$ then, by Facts 1 and 2, $K > 5$ and $\#C_j = 1$ for $j = 2, 3, 4, .., K$. Thus, merging C_K and C_{K-1} improves the ranking of all the members of the new coalition. Hence, $\#C_2 \geq 2$. If $\#C_1 = 2$, then merging C_2 with one element of C_K improves the ranking of all the members of the new coalition, who have no incentives to divide. $\blacksquare$

Since $N \geq 7$, there are at least two agents who do not belong to either C_1 or C_2. Adding either one or two of these agents to C_2, one can improve the ranking of the all members of the new coalition and create no incentives for the agents in this coalition to split. QED

Note that for $N = 6$, the partition $\{1, 5, 6\}, \{3, 4\}, \{2\}$ does not fail coalition proofness for some homogeneous Π that satisfies MW.

6 Strong stability

In this section, we refine the stability notion by requiring that social orders are stable for any coalitional power relation consistent with the individual power relation.

Definition: A social order $\Sigma = \{C_1, .., C_K\}$ is *strongly stable* if $\mathcal{S}^{\Pi}(\Sigma) = \varnothing$ for any $P-$consistent coalitional power relation Π.

Theorem 5 shows that the social order Σ^* is strongly stable. Consider the class of social orders $\mathcal{F}$ derived by modifying Σ^* recursively in the following fashion. A social order $\Sigma = \{C_1, .., C_K\}$ is in $\mathcal{F}$ if and only if

1. $C_1 = C_1^*$ or $C_1 = C_1^* \cup \{N\}$
2. $C_k = \{C_k^* / \cup_{j=1}^{k-1} C_j\}$ or $C_k = \{C_k^* / \cup_{j=1}^{k-1} C_j\} \cup \{\max(I / \cup_{j=1}^{k-1} C_j)\}$

For $N = 8$, the social order $\{\{1,3,5,7,8\},\{2,6\},\{4\}\}$ is in $\mathcal{F}$ as it is obtained by adding agent 8 to C_1^*. The reader can verify that a proof identical to the proof of Theorem 5 shows that any social order in $\mathcal{F}$ is strongly stable.

Theorem 8 *A social order Σ is a strongly stable social order if and only if $\Sigma \in \mathcal{F}$.*

Proof: We only need to show that if Σ is strongly stable then $\Sigma \in \mathcal{F}$. First, we show that a ranking consistent with a strongly stable social order must rank the agents in C_1^* as in Σ^*.

Agent 1 needs to be ranked first: consider a coalitional power relation such that $\{1\}\Pi((I/\{1\})$. To see that agent 3 must be at least second, consider Π such that $\{2,3\}\Pi(I/\{2,3\})$. Since agent 1 must be first, if agent 3 is not second, he can deviate forming a coalition with agent 2. Now consider agent 5 and assume he is ranked below the third position. Choose Π such that $\{2,4,5\}\Pi(I/\{2,4,5\})$. Since agents 1 and 3 are first and second, agents 2 and 4 can form a coalition with agent 5 and improve their position. Suppose we have shown that all agents $2i-1 \in C_1^*$ are in the i^{th} position. Consider the agent $2i+1 \in C_1^*$ and suppose he is below the $(i+1)^{th}$ position. Choose Π such that $\{2,4,6,..2i,2i+1\}\Pi(I/\{2,4,6,..2i,2i+1\})$. Since $\{1,3,5,..2i-1\}$ are ranked in the $1,...,i^{th}$ positions, agents $\{2,4,6,..2i\}$ can form a coalition with $2i+1$ and improve their position.

Now it is easy to verify that no agent in C_1^* can belong to a coalition that contains agents that are not in C_1^* and are not agent N . To show that all agents in C_1^* must belong to the same coalition, choose Π such that $\{2\}\Pi(I/\{1,2\})$.

To characterise C_2, repeat the arguments above for social orders where the set of agents is $I - C_1$. To ensure that deviations analogous to the ones in the above paragraphs are durable for when the set of agents is I, it is sufficient to consider Π's such that $\{1\}\Pi(I/\{1\})$ as no agent in C_1 would then join a counter deviation. Repeating these arguments for all C_i's concludes the proof. QED

This result is of theoretical interest in and of itself. The requirements for a strongly stable social order are severe but can be partially justified on robustness grounds. It is natural to think of the power of coalitions as more variable and harder to assess than the power of individuals. The aggregate strength of a group can depend on characteristics of social interaction that are unobservable and difficult to evaluate.

The concept of strong stability can be illustrated axiomatically in a simple and intuitive fashion. Consider the following two criteria:

(K1) A social order $\Sigma = \{C_1, .., C_K\}$ is such that C_i dominates $\cup_{j=i+1}^{K} C_j$ for any $i = 1, ..., K$.

(K2) A social order $\Sigma = \{C_1, .., C_K\}$ is such that for any agent $j \in C_i$, his rank within C_i is weakly better than his rank within the set $(\cup_{l=i+1}^{K} C_l) \cup \{j\}$.

(K1) is a criterion for external stability: all coalitions are immune from the threat of a unified challenge coming from all weaker coalitions. (K2) is a criterion for internal stability in that agents in a coalition never wish to join a united challenge by all weaker coalitions.

Proposition 9 *Σ satisfies (K1) and (K2) if and only if $\Sigma \in \mathcal{F}$.*

Proof: First note that any $\Sigma \in \mathcal{F}$ satisfies (K1) and (K2). Now consider a social order $\Sigma = \{C_1, .., C_K\}$ that satisfies (K1) and (K2). By (K1), agent 1 is in C_1. By (K2), agent 2 cannot be in C_1. By (K1) again, agent 3 is in C_1. By (K2) again, agent 4 cannot be in C_1. Repeating these arguments up to $\max C_1$ implies that $C_1 \supset C_1^*$. Recall that $\# C_1^* \geq \dfrac{N}{2}$. If $C_1 / C_1^* \neq \phi$, (K2) implies that any agent $j \in C_1/C_1^*$ must be ranked worse than all agents in $\cup_{l=i+1}^{K} C_l$ and $\# C_1/C_1^* \leq 1$. The same arguments for the other coalitions establish that $\Sigma \in \mathcal{F}$. QED

Strongly stable social orders depend critically on coalition leaders. In particular, eliminating agent 1 from the set of agents causes a major upset in the social structure. The new most powerful coalition will be composed of agents that were not in C_1^*, while those agents who were in C_1^* are now divided into smaller and less powerful coalitions. In contrast, eliminating the lowest individual in society doesn't affect the social order except for the absence of that agent.

The social order Σ^* can be obtained axiomatically by strengthening criterion (K2) . Consider

(K3) A social order $\Sigma = \{C_1, .., C_K\}$ is such that $C_i / \min\{C_i\}$ is dominated by $\cup_{j=i+1}^{k} C_j$ for any $i = 1, ..., K-1$ and $\min\{C_K\}$ dominates $C_K / \min\{C_K\}$.

Proposition 10 Σ^* *is the unique social order which satisfies (K1) and (K3).*

Proof: First note that all coalitions of Σ^* satisfy (K1) and (K3). Now consider a social order $\Sigma = \{C_1, .., C_K\}$. First note that by (K1) and (K3), if N is even, $\#C_1 = \frac{N}{2}$, if odd, $C_1 = \frac{N+1}{2}$. By (K1), agent 1 is in C_1. By (K3) agent 2 cannot be in C_1. Suppose agent 3 is not in C_1 then (K1) is contradicted as C_1 cannot dominate $\cup_{j=1+1}^{k} C_j$. Repeating these arguments will imply that $C_1 = C_1^*$. Repeating the same arguments for the other coalitions establishes that $C_i = C_i^*$ for $i = 2, ..., K$. QED

7 Special power relations

We are unable to provide a complete characterisation of stable social orders under arbitrary power relation. In this section, we explore social orders that are stable for particular power relations.

7.1 Congruence

Generally, stable social orders yield a power ranking that differ from the exogenous individual power relation. In the next proposition we give necessary and sufficient conditions for the societal power relation induced by the social order to be identical to the individual power relation.

Proposition 11 *There exists a stable social order* Σ *such that* $Q^{\Pi}(\Sigma) = P$ *if and only if* $\{1\}\Pi\{2, 3, ..., N\}$.

Proof: If $\{1\}\Pi\{2,3,...,N\}$, consider the social order with only one coalition. For any any deviating coalition C, let j be the most powerful agent in C. It is easy to verify that j's position cannot be better than the j^{th} position. Consider a social order Σ such that $Q^{\Pi}(\Sigma) = P$ and assume that $\{2,3,...,N\}\Pi\{1\}$. Obviously $\{2,3,...,N\}$ is a durable deviation and therefore Σ is not stable. QED

7.2 Homogeneous Power

Suppose that the power of agents is approximately the same. The following result shows that, in a stable social order, the most powerful coalition must exclude some of the most powerful agents.

Proposition 12 *Consider a stable social order* Σ*. If* Π *is homogeneous, it is impossible that* $\{1,2,3\} \subset C_1$.

Proof: Consider first the case of N even and suppose that there exist a stable Σ such that $\{1,2,3\} \subset C_1$. To obtain a contradiction, take a coalition with agents 2, 3, and all agents ranked strictly lower than $\dfrac{N}{2}+1$ in Σ. This deviation is durable. Now consider the case of N even and suppose that there exist a stable Σ such that $\{1,2,3\} \subset C_1$. Take a coalition with agent 2 and all agents ranked strictly lower than $\dfrac{N+1}{2}$ in Σ. This deviation is again durable. QED

We say that Π is *quasi-lexicographic* if $A\Pi B$ whenever $\#A = \#B$ and $\min A < \min B$. We now show that, when the distribution of power is homogenous and quasi-lexicographic, two criteria are sufficient to yield stable social orders. These criteria are (K3) and a weakening of (K1) that we denote by (K1').

(K1') A social order $\Sigma = \{C_1, .., C_K\}$ is such that $C_i \Pi \cup_{j=i+1}^{K} C_j$ for any $i = 1, ..., K$.

Proposition 13 *Suppose that Π is homogeneous and quasi lexicographic. A social order Σ that satisfies (K1') and (K3) is stable.*

Proof:

Let $\Sigma = \{C_1, ..., C_K\}$. Let $\#C_1 = L$. Note that (K1'), (K3), and homogeneity of Π imply that:

(1) if N is odd, then $\#(\bigcup_{i=2}^{K} C_i) = L - 1 = \dfrac{N-1}{2}$;

(2) if N is even then $\#(\bigcup_{i=2}^{K} C_i) = L = \dfrac{N}{2}$ and $\min C_1 = 1$.

We first show that for any coalition C such that $V_i^{\Pi}(\Sigma \top C) < V_i^{\Pi}(\Sigma)$ for any $i \in C$ and $C_1 \cap C \neq \varnothing$, there exists a durable counter-deviation.

Fact 1: $\#C < L$.

Proof: Suppose that $\#C \geqslant L$. Note that $\min C_1 \notin C$ and that, by (K3), $C_1 / \min C_1$ is dominated by $\bigcup_{i=2}^{K} C_i$. Suppose that $\#(C \cap C_1) = R$. Then, there are at least R agents in $\bigcup_{i=2}^{k} C_i$ that are not in C. To see this, suppose that $\min(C \cap C_1)$ is in the m^{th} position in Σ. By (K3), there are $m-1$ agents in $\bigcup_{i=2}^{K} C_i$ who are more powerful than $\min(C \cap C_1)$. Denote the set of these agents by D and let $E = D \cap C$. Since the position of $\min(C \cap C_1)$ in $\Sigma \top C$ is worse than or equal to $\#E + R$, we need $\#E + R < m$. Thus, $\#D - \#E > m - 1 - m + R = R - 1$, implying that at least R agents from $\bigcup_{i=2}^{K} C_i$ are not in C.

Hence, $\#C \leq \#(\bigcup_{i=2}^{k} C_i)$. If N is odd, the claim follows by (1). If N is even and $\#C = L$, then, as Π is homogeneous and is quasi lexicographic, $\bigcup_{i=1}^{K} C_i / C$ is a durable deviation as it includes agent 1 by (2). $\blacksquare$

Fact 2: $C \cap (N/C_1) \neq \phi$.

Proof: Suppose that $C \cap (N/C_1) = \phi$. Then $\#C > \dfrac{L}{2}$ and $\#(C_1/C) < \dfrac{L}{2}$. If N is even, then $\min C_1 = 1$ Agent 1 and all the $\dfrac{N}{2} - 1$ worst ranked agents in $\Sigma \top C$ form a durable counter-deviation. If N is odd and L is even, $\#C_2 = \dfrac{L}{2}$ and so, in $\Sigma \top C$, C is ranked first and C_2 second. Since $\#(C \cup C_2) > L$ and, by Fact 1, $\#C < L$, the agents ranked $\left(\dfrac{N+1}{2}\right)^{th}$ and $\left(\dfrac{N+3}{2}\right)^{th}$ are both in C_2. So taking all $\dfrac{N+1}{2}$ worst ranked agents in $\Sigma \top C$ will be a durable counter-deviation since the agent ranked $\left(\dfrac{N+1}{2}\right)^{th}$ in $\Sigma \top C$ strictly improves his position.

Finally, assume that N is odd and L is odd. If in $\Sigma \top C$ C is first, C_1/C is second, and C_2 third, we have that $\#(C_1/C) = \#C_2 = \dfrac{L-1}{2}$ and so $\min C_1 < \min C_2$. Therefore, $\min C_1$ is more powerful than at least $\dfrac{L-1}{2}$ agents in $\bigcup_{i=2}^{k} C_i$. Since $\#C > \dfrac{L}{2}$, $\min C_1$ is ranked worse than or at the $\left(\dfrac{L+3}{2}\right)^{th}$ position in $\Sigma \top C$. Therefore, a coalition consisting of $\min C_1$ and the $L-1$ worst ranked agents in $\Sigma \top C$ is a durable deviation since $\min C_1$ will be ranked better than or at the $\left(\dfrac{L+1}{2}\right)^{th}$ position. ■

We now consider the following cases:

Case 1 C is first and C_1/C is second.

By Facts 1 and 2, $\#((C_1/C) \cup C) > L$. If N is odd, take a coalition C' containing all the L worst ranked agents in $\Sigma \top C$. The coalition C' is a durable deviation as the L^{th} and the $(L+1)^{th}$-ranked agents in $\Sigma \top C$ are both in C_1/C and hence, the L^{th}-ranked agent is ranked strictly better than the L^{th} position in $(\Sigma \top C) \top C'$. If N is even, then $\min C_1 = 1$ and agent 1 together with the $L-1$ worst ranked in $\Sigma \top C$ can form a durable deviation. ■

Case 2 C is first in $\Sigma \top C$ and C_1/C is worse that second.

Let the second ranked coalition in $\Sigma \intercal C$ be denoted by $\bar{C}$. By homogeneity, $\#\bar{C} \geq \#((C_1/C)$ and, by Facts 1 and 2, $\#(\bar{C} \cup C) > L$. If N is odd, a coalition of the L worst ranked agents in $\Sigma \intercal C$ is a durable deviation for the same reasons as in Case 1. If N is even, take the L worst ranked agents in $\Sigma \intercal C$ to form a durable deviation as $\min C_1 = 1$ among these agents. ■

Case 3 C is not ranked first.

Let $C_1^{\Sigma \intercal C}$ be the best ranked coalition in $\Sigma \intercal C$. Then, $\#C_1^{\Sigma \intercal C} \geqslant 2$ and thus the agents ranked first, second and third in C_1 are not in C. This implies that $\#C_1^{\Sigma \intercal C} \geqslant 3$. Continuing in this fashion, a contradiction is obtained when we conclude that $C_1 \cap C = \varnothing$. ■

We now prove the Proposition. Consider any C such that $C_1 \cap C = \varnothing$, $C_2 \cap C \neq \varnothing$, and $V_i^{\Pi}(\Sigma \intercal C) < V_i^{\Pi}(\Sigma)$ for each $i \in C$. By repeating the arguments above (ignoring the agents in C_1) we can construct a counter-deviation C'' to C in $\Sigma \intercal C$ that satisfies $C''\Pi((\cup_{i=2}^{K} C_i)/C'')$. Note that C'' is durable; any counter-deviation to C'' in $(\Sigma \intercal C) \intercal C''$ must include an agent in C_1 and, by the above arguments, will itself have a durable counter-deviation.

The completion of the proof is obtained by an inductive repetition of the above arguments. QED

As an application of Proposition 13, we show the existence of stable social orders in which coalitions are extremely heterogeneous, and are formed by a powerful agent and some of the weakest available agents. Define the algorithm that determines the allocation $\tilde{\Sigma} = \{\tilde{C}_1, ..., \tilde{C}_K\}$ as follows. Let $\tilde{C}_1 = \{1, j, j+1, ..., N\}$ where j is chosen so that $0 \leq \#\tilde{C}_1 - \#(N/\tilde{C}_1) \leq 1$. Having defined $\tilde{C}_k$ for $k \leq k'$, define $\tilde{C}_{k'+1} = \{k'+1, j, j+1, ..., \max(N/\cup_{i=1}^{k'} \tilde{C}_i)\}$, where j is chosen so that $0 \leq \#\tilde{C}_{k'+1} - \#(N/\cup_{i=1}^{k'+1} \tilde{C}_1) \leq 1$.

Corollary 14 *Suppose that* Π *is homogeneous and is quasi-lexicographic. Then* $\tilde{\Sigma}$ *is stable.*

Proof: It is straightforward to verify that $\tilde{\Sigma}$ satisfies (K1') and (K3).

References

[1] Acemoglu D, G Egorov and K Sonin (2007), "Coalition Formation in Non-Democracies", working paper.

[2] Aumann R and M Maschler (1964), "The Bargaining Set for Cooperative Games", in *Advances in Games Theory* (M Dresher, LS Shapley and AW Tucker eds), Princeton University Press, 443-476.

[3] Chwe M (1994), "Farsighted Coalitional Stability" *Journal of Economic Theory,* 299-325.

[4] Diamantoudi E and L Xue (2007), Coalitions, Agreements and Efficiency", *Journal of Economic Theory,* 105-125.

[5] Greenberg J, *The Theory of Social Situations*, Cambridge University Press, 1990.

[6] Jordan J (2006a), "Pillage and Property", *Journal of Economic Theory*, 26-44.

[7] Jordan J (2006b), "Power and Legitimacy in Pillage Games", working paper.

[8] Lucas W and R Thrall (1963), "N-Person Games in Partition Function Form", *Naval Research Logistics Quarterly*, 281-298.

[9] Myerson R (1977), "Values of Games in Partition Function Form", *International Journal of Game Theory*, 23-31.

[10] Myerson R, *Game Theory*, Harvard University Press, 1991.

[11] Osborne M and A Rubinstein, *A Course in Game Theory*, MIT Press, 1994.

[12] Piccione M and A Rubinstein (2004), "Equilibrium in the Jungle", working paper.

[13] Piccione M and A Rubinstein (2007), "Equilibrium in the Jungle", *Economic Journal*, 883-896.

[14] Ray D and R Vohra (1997), "Equilibrium Binding Agreements", *Journal of Economic Theory,* 30-78.

[15] Ray D and R Vohra (1999), "A Theory of Endogenous Coalition Structure", *Games and Economic Behavior*, 286–336.

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