Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Internal Hierarchy and Stable Coalition Structures

Massimo Morelli and In-Uck Park

Working Paper n. 528

This Version: October 30, 2014

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

Internal Hierarchy and Stable Coalition Structures*

MASSIMO MORELLI
Bocconi University, IGIER
Columbia University and NBER

IN-UCK PARK
University of Bristol
Sungkyunkwan University

October 30, 2014

Abstract. *In deciding whether to join a coalition or not, an agent must consider both i) the expected power of the coalition and ii) her position in the vertical structure within the coalition. We establish the existence of a positive relationship between the degree of inequality in remuneration across ranks within coalitions and the number of coalitions to be formed endogenously in stable systems. An inherent feature of such coalitions is that they are mixed and balanced, rather than segregated, in terms of members abilities. When the surplus of a coalition is assumed to be linear in its relative power conditional on its size, we also establish the existence of stable systems and characterise them fully: a system is stable if and only if all coalitions are of an efficient size and every agent is paid her marginal contribution.* (*JEL* Codes: C71, D71)

Keywords: Stable systems, Abilities, Hierarchy, Cyclic partition.

1 Introduction

Circumstances abound in which individual agents interact via the organisations they choose to belong to. From each agent's perspective, the consequences of joining one organisation or another are determined by (i) the outcome resulting from the interaction between the organisation she chooses to join and its rival organisations, and (ii) the effect of that outcome on her within the organisation. The latter aspect is likely to be determined by the internal structure of the organisation and her position in it.

*We thank Ken Binmore, Antonio Cabrales, Jon Eguia, Piero Gottardi, Sanjeev Goyal, Johannes Hörner, Matthew O. Jackson, Hideo Konishi, David Levine, Jaime Luque, Meg Meyer, Andy Newman, Scott Page, James Peck, Chen-Yu Pan, Nicola Persico, Clara Ponsati, Debraj Ray, Daniel Seidmann, Rajiv Sethi, Paolo Siconolfi, David Strauss, Fernando Vega-Redondo, Schlomo Weber and Peyton Young for very useful suggestions. We are also grateful for comments from workshop participants at the Bristol Conference on Social Networks in 2010, the 2011 Coalition Theory Network in Barcelona, the Decentralization Conference at Ohio State and the V Workshop on Institutions, Individual Behavior and Economic Outcomes in Alghero, as well as seminar participants at the Universities of Bologna, Columbia, Madrid, Melbourne, Montreal, Oxford, Yale, Yonsei, and European University Institute. The usual disclaimer applies. Park gratefully acknowledges the Leverhulme Trust research grant.

In a political setting, for example, politicians form parties and members of each party decide on the party line and on the campaign strategy, given the perception of their strengths and the opponents' characteristics. The election outcome will be determined by the kinds of parties formed and their relative strengths. Finally, the members of the winning party will be allocated a role depending on their relative positions within the party, which will shape their payoffs. Similar descriptions apply to the world of gangs and their members and to entrepreneurial organisations, to name a social and a business example, respectively. In these situations, the agents' *ranks* within the organisation appear to be an important factor in determining their final payoffs.[1]

Understanding what determines the number and composition of coalitions (e.g., party systems, market concentration, economic and political integration) has been a recurrent focus in many strands of literature (discussed below), but, to the best of our knowledge, no systematic work has examined the relationship between such *horizontal* segmentation incentives and the *vertical* structure within each endogenous organisation. We believe, especially in contexts in which the relevant agents are heterogeneous in ability, that studying the interplay of these two dimensions could be very insightful. As ability differentials among agents increase, would there be more or less competition (in terms of the number of rival organisations to be formed) and would the organisations become more or less segregated? Are there general connections between endogenous meritocracy and the degree of competition? This paper develops a cooperative game theoretical framework to address such questions in an institution-free environment and provides some robust answers.[2]

The relevant players in our analysis are all those agents who participate in choosing which coalitions to form, merge with or split from.[3] The first key assumption of our model is that the relevant agents have heterogeneous observable abilities, and the total surplus of a coalition depends positively both on its size, and on the aggregate ability of its members, called its power. Second, we assume that each endogenous coalition has a vertical structure, in which the coalition members are ranked, or assigned to different tasks of rankable importance, and payoff shares

[1]Our model and results will relate more to the formation of *competing* parties, firms or gangs, than to the formation of clubs and jurisdictions, given that typically models on the latter are about sorting or matching preferences (for example on local public goods) and do not focus on vertical differentiation among agents.

[2]The limited literature on party formation focuses almost exclusively on the incentives that different institutional systems provide to form parties to represent different (horizontal) segments of the voters' population, whereas the impact of the internal organisation of parties on the stability of different party systems has rarely been studied (as a notable exception, see Persico et al., 2008). Intuitively, the choice between becoming the leader of a new party or remaining at a lower rank of an existing party must depend on how the different ranks are treated.

[3]In each application there could be other agents such as voters (in the political application), consumers (in the industrial organisation application) and victims (in the criminal organisations application) who do not directly engage in such decisions.

are non-decreasing in rank. Given a distribution of agents' abilities and a surplus function for all possible coalitions, in line with standard core-stability, we define a system – a partition of agents into coalitions and an imputation rule of how to share the surplus among members for each coalition – as being *stable* if there is no deviating coalition where every member fares better than in the payoff distribution of the original system.

The most robust result of our analysis is that the more equally shared is the surplus across ranks internally, i.e., the lower is the vertical inequality, the smaller is the number of rival organisations that can be sustained in a stable partition of the relevant agents. To be more precise, focusing attention on "symmetric" stable systems in which all coalitions within each system adopt the same imputation rule of sharing their surplus across members of different ranks, Theorem 1 establishes that if one symmetric stable system exhibits a more equal sharing rule than another, then the former system consists of fewer coalitions. This result is then extended to non-symmetric stable systems and to environments with finite but large populations (under some technical assumptions).

When the surplus of a coalition is assumed to be linear in its power (i.e., when the power and the size of the coalition enter multiplicatively in the surplus function), we also establish the existence of stable systems and fully characterise them: A system is stable if and only if all coalitions are of an efficient size and every agent is paid her marginal contribution (unless the population size falls short of the efficient coalition size, in which case a suitably modified result is obtained). We then consider a heuristic subclass of environments in which agents' ability levels are distributed geometrically, and show that a system is stable if and only if the agents are partitioned into *cyclic* coalitions which are coalitions composed of equidistant agents in their ability ordering.

Our paper highlights several features of the endogenous formation of rival organisations that are novel relative to the existing literature on coalition formation: (1) the more unequal is the allocation of payoffs, the more fragmented will be the rival organisations to be formed; (2) organisations of different internal norms may coexist; and (3) organisations tend to consist of members from widely dispersed ability levels. The last feature, in particular, contrasts starkly with the segregation outcomes that are prevalent in the literature on some other types of group formation, such as the important literature on clubs and jurisdictions providing local public goods, preluded by Tiebout (1956).[4]

[4] A chronological sequence of formal advancements in that literature can be found in Ellickson (1973), Westhoff (1977), Wooders (1978), Guesnerie and Oddou (1981), Greenberg and Weber (1986), Demange (1994), Konishi, Le Breton and Weber (1998), Ellickson et al (1999), Casella (2001), Jehiel and Scotchmer (2001), and Zame (2009). These studies differ from ours in that different jurisdictions provide different local public good quantities and the endogenous coalitions do not play a constant sum game. Moreover, typically agents are not differentiated in terms of ability.

The literature on social classes (Akerlof, 1997), partnerships (Farrell and Scotchmer, 1988), hedonic games (Banerjee, *et al.,* 2001; Bogomolnaia and Jackson, 2002; Le Breton, et al., 2008; Watts, 2007), social status (Milchtaich and Winter, 2002), and organisation (e.g., Demange 2004; Garicano and Rossi-Hansberg, 2006; and an earlier work on firm formation by Legros and Newman, 1996), is all related in a broad sense to what we do, but our approach is distinguished from these studies in the following respect: agents of *vertically differentiated abilities* compete to form *endogenous* coalitions with attractive counterparts through a mutually favourable surplus sharing rule. Thus, agents potentially face a dilemma between teaming up with more able people for a more powerful coalition and teaming up with less able people for a higher internal rank. Damiano, *et al.* (2010) consider a similar tension but in a setting where agents decide which one to join from a *fixed* set of coalitions, motivated by contexts different from ours.[5] Watts (2007), on the other hand, analyzes two separate settings, one in which agents desire to team up with higher ability members (under the "average quality payoff"), and an opposite one in which they desire to team up with lower ability members (under the "big fish payoff").

Piccione and Razin (2011) study coalition formation in partition function games, in which an agent's social ranking is determined lexicographically, first by the "power relation" between the coalitions formed, then by her ability within the relevant coalition. The core is empty in this setting if the size dictates the power relation of coalitions. For this reason, they define a recursively stable solution concept, yielding existence and characterisation results in the spirit of our non-segregation results. A more recent paper, Barbera, *et al.* (2014), studies a particular case of our linear surplus function, with size entering as a step function. In this context they show that meritocratic sharing norms in some coalitions can coexist with egalitarian norms in others. Two additional differences relative to this work are that we allow for all possible sharing rules (in addition to the two they focus on) and that we do not impose any tie-breaking rule. Our analysis shows that the only sharing rule that survives in this case is the meritocratic one (cf. footnote 13).

The current paper also makes a conceptual contribution to political economy, and in particular to the literature on party formation, showing that even with similar institutions and preferences, different party systems can be stable, depending on the parties' internal organisations.[6] For more distantly related work on trade alliance

[5] In Damiano, *et al.* (2010), agents of different abilities choose between two organisations of a fixed capacity of measure 1, when the agent's utility increases both in the average ability of the organisation (peer effect) and in her internal ranking (pecking order effect). If the value of each coalition is a function of the average ability of agents, they obtain some degree of segregation of ability types, with a larger overlap when the pecking order effect is stronger. Their results apply to very different contexts, such as students' choices among a fixed set of universities, rather than endogenous formation of organisations.

[6] On the importance of party formation and pre-election coalition formation across systems, see, e.g., Levy (2004), Morelli (2004) and Bandyopadhyay, Chatterjee and Sjöström (2010). See also Dhillon (2005) for a survey.

formation, see, e.g., Yi (1996) and Casella and Feinstein (2002).

We adopt a cooperative approach in which individual agents can coordinate on a deviation as a group.[7] Hence, parties, firms, teams or gangs are more natural types of coalitions that fit our analysis than are countries/jurisdictions where agents individually decide whether to move in or out, such as in Jehiel and Scotchmer (2001).

In hedonic games coalitional deviations are allowed, but agents' payoffs are determined by the composition of their own coalition only. In our game the agents' utility depends on the rank and the degree of vertical inequality in the coalition, as well as on the aggregate strength of the coalition, so it is not a proper hedonic game.[8] Our model can also be viewed as generalising Gamson games (see, e.g., Le Breton, et al., 2008): in this special class of hedonic games the total cake goes to the coalition that has more than half of the total talent, whereas our analysis includes settings where coalitions fight over market shares or power shares, with no magic value given to passing a 50% threshold.

The paper is organised as follows. Section 2 describes the general model, specifying the class of environments for which our general results can be proved. Section 3 establishes as generally as possible that if a stable system exhibits a higher level of vertical inequality than another, then the former must consist of a (weakly) larger number of competing coalitions than the latter. Section 4 focuses attention on environments in which a coalition's surplus is linear in its power, confirms the existence and provides a full characterisation of stable systems, and then presents illustrations of all stable systems for the special case that agents' abilities are geometrically distributed and coalitions choose a single parameter imputation rule that captures the common payoff inequality between any two adjacent ranks. Section 5 offers concluding remarks.

2 Model

Consider an economy with a large population $\Omega = \{1, 2, \cdots, N\}$ consisting of N agents of heterogeneous ability. For expositional ease, we conduct our analysis primarily for the limit case of $N \to \infty$, i.e., with countably infinite agents or $\Omega = \mathbb{N}$, but also discuss how the results are extended to the case of a large but finite number of agents. Each agent $i \in \Omega$ has an observable ability $a_i > 0$ (which could be political ability, market ability, or criminal ability, etc., depending on the

[7] See, e.g., Aumann and Drèze (1974) for an early study on the cooperative stability of coalition structures.

[8] When we fix the degree of vertical inequality and impose the ability ranking assumption, our game is hedonic. However, generally it does not satisfy known conditions for existence of a core, namely, balancedness (Scarf, 1967) and top-coalition property (Banerjee et al, 2001), neither in a finite setting nor when naturally extended to an infinite setting.

application). Agents are indexed according to their ability, with the convention that $a_i > a_{i+1}$ for all i. (The main results straightforwardly extend to the case that $a_i \geq a_{i+1}$ at a cost of expositional complication, the details of which are omitted here.) We assume that the agents' abilities add up to a finite number, which we normalise to 1, i.e., $\sum_{i\in\Omega} a_i = 1$.

We postulate that the total surplus $S(Z)$ of a coalition $Z \subset \Omega$ of agents depends both on the "coalitional productivity/power", $p(Z) = \sum_{i\in Z} a_i$, and the size/cardinality of the coalition, $|Z|$:

$$S(Z) = s(p(Z), |Z|)$$

where $s : [0,1] \times \mathbb{R} \to \mathbb{R}$ is a C^2 function such that

$$\underline{v} < s_1(p,q) < \bar{v}, \;\; s_2(p,q) > 0, \;\; s_{22}(p,q) < 0, \;\; s(0,q) = 0 \;\; \text{and} \;\; \lim_{q\to\infty} s(1,q) < \infty \quad (1)$$

with $0 < \underline{v} < \bar{v} < \infty$, where s_j represents the partial derivative of s with respect to the j-th argument for $j = 1, 2$, and s_{22} the relevant second partial derivative. Hence, the positive effect of a coalition's power on its surplus is bounded below and above; the effect of size is positive but concave and vanishes eventually. We extend s to coalitions of an infinite size by defining $s(p,\infty) = \lim_{q\to\infty} s(p,q)$ for all $p \in [0,1]$.

A population, Ω, the agents' ability levels, and a surplus function s, specify an "environment" in which agents may form coalitions endogenously and share surpluses within coalitions. Below we define a stable outcome of such a process.

Each coalition being formed, $Z \subset \Omega$, will adopt an *imputation rule,* denoted by $f : Z \to [0,1]$, that specifies for each member $i \in Z$ a fraction $f(i)$ of the total surplus, $S(Z)$, to be allocated to that member. Note that $\sum_{i\in Z} f(i) = 1$. Thus, the payoff of agent $i \in Z$ is

$$u_i(Z,f) \;=\; f(i) \cdot S(Z). \quad (2)$$

A *system* is a pair (π, ρ) where $\pi = \{Z_1, \cdots, Z_K\}$ is a partition of agents in Ω into coalitions and ρ is a function that maps each coalition $Z_k \in \pi$ to an imputation rule of that coalition. We adopt the convention that the coalitions are labelled according to the order of ability of the most able member of each coalition. That is, $\min\{i|i \in Z_k\} < \min\{i|i \in Z_{k'}\}$ if $k < k'$.

A system (π, ρ) is *stable* if there does not exist a deviation $D \subset \Omega$ that is *profitable* relative to the system (π, ρ) in the sense that

$$u_i(D,f) \;\geq\; u_i(\pi(i), \rho(\pi(i))) \quad \forall i \in D \neq \emptyset$$

for some imputation rule f that D may adopt, where $\pi(i)$ is the coalition $Z_k \in \pi$ such that $i \in Z_k$, and the inequality is strict for some $i \in D$.

In terms of interpretation, the coalition's power, $p(Z)$, and its size, $|Z|$, represent the quality and quantity, respectively, of the human resources available within an organisation for its operation. For an organisation to function properly, typically various distinct tasks need to be performed. The property that the surplus increases in $|Z|$ as well as in $p(Z)$ reflects that not only these tasks are better performed by more able agents, but there are also some benefits from specialisation (although they diminish as the coalition grows in size). As such, an agent's contribution to a coalition depends on its size and power. In a stable system, as we will see, every agent's compensation should be commensurate with what she contributes to the coalition's total surplus, for there to be no profitable deviation.

We are interested in whether and how the endogenous level of inequality in compensation within a coalition is related to the structure of endogenous coalitions (also called organisations for their implicit assignment of tasks to agents of different ability levels). To facilitate exposition, we introduce the notion of the "rank" of members of an organisation according to their compensation, i.e., the agent who gets paid the most within an organisation is ranked first, and so on. This is consistent with the usual notion of rank, provided that agents are paid in line with their contributions and high rank individuals are responsible for more important tasks with greater expected impacts. Arbitrary ranking is allowed among the agents who get paid the same within an organisation, as it is inconsequential in our analysis.

Note from (1) that a more able member of an organisation makes a larger contribution to the total surplus. Consequently, it is intuitively appealing that members' ranks in compensation coincide with their ranks in ability within an organisation. Indeed, it is innocuous to restrict our attention to stable systems with such a property according to the next lemma, for the purpose of characterising the number and composition of coalitions in stable systems. The lemma establishes that, even if relative payoffs are not tied to relative abilities explicitly, they are tied endogenously due to competition among rival coalitions. We note that this result holds regardless of the cardinality of Ω.

Lemma 1 *Suppose a system (π, ρ) is stable and the imputation rule f of a coalition $Z \in \pi$ ranks its members differently from their ability ranking. The system continues to be stable when the imputation rule of Z is modified in such a way that the agent whose ability rank is $r \in \mathbb{N}$ in Z gets the r-th highest compensation according to f, that is, the agent whose ability rank is r in Z gets*

$$f_r \cdot S(Z) \quad \text{where} \quad f_r = \max_x \{x \in [0,1] \,|\, \#\{i \in Z : f(i) \geq x\} \geq r\}.$$

Proof. Suppose that there is a stable system (π, ρ) described in the lemma. Consider an arbitrary pair of agents in Z, say i and j where $i < j$, but j is ranked above i in terms of compensation. This can happen in one of the following two ways.

First, if $f(i) = f(j)$ yet j is ranked above i because any ranking is allowed among equally paid members, then simply reverse their rankings. Then, the modified system is stable because every agent gets the same utility in both systems.

Second, if $f(i) < f(j)$ then modify f by swapping $f(i)$ and $f(j)$, so that i is paid more than j in the "new system." To verify that the new system is stable by way of contradiction, suppose to the contrary that it is not stable. Then, a profitable deviation D, with a deviation imputation rule f^D, exists that includes either i or j. If D included only i, then the same D with f^D would be profitable relative to the original system as well, because i is paid less in the original system than in the new system. If D included only j, then D and f^D, with j replaced by i, would be profitable relative to the original system, because i in the original system is paid the same as j in the new system (and $S(D)$ increases when $j \in D$ is replaced by i). If D included both i and j, then the same D and f^D with $f^D(i)$ and $f^D(j)$ swapped, would be profitable relative to the original system, because i (j) in the original system is paid the same as j (i) in the new system.

By sequentially swapping $f(i)$ and $f(j)$ if $i < j$ but $f(i) < f(j)$ for members i and j in Z, we can construct a new stable system with the same partition π where the agent whose ability rank is $r \in \mathbb{N}$ in Z gets the r-th highest compensation according to its imputation rule, thereby completing the proof.[9] ■

In light of this lemma, we take it for granted in what follows that any stable system satisfies the property that the internal ranking of members induced by the imputation rule coincides with their ability ordering within the organisation, which we refer to as the "ability ranking" property. Hence, we may represent an imputation rule of a coalition $Z \subset \Omega$ by a vector

$$\mathbf{f} = (f_1, f_2, \cdots) \in [0,1]^{|Z|}$$

where f_r is the fraction of the surplus rewarded to the agent, say $i \in Z$, who is ranked r-th in Z according to ability. Note that $\sum_{r=1}^{|Z|} f_r = 1$.

We measure inequality within an organisation by the ratios of the payoff each rank gets relative to that one rank above gets, i.e., f_{r+1}/f_r for $r = 1, 2, \cdots$.

3 General Results

Our core results concern the relationship between internal inequality and an endogenous stable coalition structure, in particular, the number of coalitions that emerge and their compositions. It proves useful to start with the cases in which the degrees of internal inequality can be compared easily between systems. Specifically, first we consider "symmetric systems" in which the imputation rules of all coalitions within

[9] To be fully precise, there is a technical complication due to the possibility that one may not finish the swapping process in finite steps when there are infinite instances of "reversed" ranking initially. However, this proof is sufficient for our results because the proofs of Theorems 1 and 2 only require coincidence of compensation ranking and ability ordering for the M most able agents for some large $M < \infty$.

a system exhibit identical levels of inequality in the following sense: for each rank $r = 1, 2, \cdots$, the ratio f_{r+1}/f_r is the same across all coalitions whenever well-defined. Note that the actual values of f_r differ across coalitions of different sizes. Nonetheless, we can denote the imputation rule component of a symmetric system (π, ρ) simply by the imputation rule of the coalition of the largest size, say $\mathbf{f}$, with the interpretation that for a coalition $Z_k \in \pi$ of a smaller size, say L, the imputation rule is the truncation of $\mathbf{f}$ up to the first L entries as fractions to the sum of them. Later, we explain how our results are extended beyond symmetric systems.

Fix an environment, i.e., a population Ω, a specification of agents' abilities, and a surplus function s. Consider two symmetric systems, denoted by $F = (\pi_F, \rho_F)$ and $G = (\pi_G, \rho_G)$. Let $\mathbf{f} = (f_1, f_2, \cdots)$ and $\mathbf{g} = (g_1, g_2, \cdots)$ denote, respectively, the imputation rule of the largest size coalition in F and G. We say that F is *less equal* than G if

$$\frac{f_{r+1}}{f_r} \leq \frac{g_{r+1}}{g_r} \quad \forall r = 1, 2 \cdots \tag{3}$$

whenever both ratios are well-defined, with at least one strict inequality. We now state our first general result. For brevity, we use SSS as the acronym for "stable symmetric system" in the sequel.

Theorem 1 *Suppose there are two SSS's in an environment with $|\Omega| = \infty$ and*

$$\limsup_{n \to \infty} \frac{a_{n+1}}{a_n} < 1. \tag{4}$$

If one system is less equal than the other, then the former consists of a weakly larger number of coalitions than the latter.

Even though there exist ability distributions with a finite sum that fail the condition (4), it is satisfied by a large class of ability distributions, including geometric distributions, i.e., $a_n = \alpha^n$ for any $\alpha \in (0, 1)$. In addition, (4) is a sufficient (rather than necessary) condition, so the result may hold even when it is not satisfied.

Here we briefly outline the basic arguments underlying this result (and defer a formal proof to the Appendix). Consider two stable symmetric systems, F and G, the former less equal than the latter. For each rank r, consider the truncated coalitions of F and G consisting of those members who are originally ranked r or below in their respective coalitions. By (3), as r gets large, the fractions of the total surplus that the members of these truncations collectively receive in the original coalitions dwindle more rapidly in the less equal system, F. In addition, what they receive collectively is in line with what they contribute collectively, for otherwise there would be an incentive either for the original coalition to sever the truncation or for the truncation to deviate by separating out. Since what they contribute collectively is essentially proportional to the sum of their abilities, ability ranking implies that this sum shrinks more rapidly in F as r increases than it does in G.

This means that ability level tends to shrink more rapidly as descending down the rank in coalitions of F, which would not be possible if there were fewer coalitions in F than in G. The logic outlined above carries a further implication that non-segregation is an inherent feature of stable systems, that is, rival coalitions consist of agents whose ability levels are interspersed over a wide range.

This logic does not hinge on the condition that F and G are symmetric systems. Therefore, it is straightforward to verify that the same result extends to non-symmetric stable systems, $F = (\pi_F, \rho_F)$ and $G = (\pi_G, \rho_G)$, by modifying the definition of being a less equal system as follows: Let $(f_1^k, f_2^k, \cdots) = \rho_F(Y_k)$ be the imputation rule of a coalition $Y_k \in \pi_F$, and let $(g_1^\ell, g_2^\ell, \cdots) = \rho_G(Z_\ell)$ be the imputation rule of a coalition $Z_\ell \in \pi_G$. Then, F is less equal than G if

$$\max_{\{k|Y_k \in \pi_F\}} \frac{f_{r+1}^k}{f_r^k} \le \min_{\{\ell|Z_\ell \in \pi_G\}} \frac{g_{r+1}^\ell}{g_r^\ell} \quad \forall r = 1, 2 \cdots \tag{5}$$

So far, for analytic ease, we have considered a countably infinite population of agents ($|\Omega| = \infty$) and established a positive relationship between internal inequality and the number of coalitions to be formed. We now show that this relationship continues to hold when there are a large but finite number of agents under some technical conditions, which basically stipulate that the positive power effect on the surplus increases for coalitions of sufficiently small power ((6) below), whereas the positive size effect on the surplus vanishes sufficiently rapidly for coalitions of large sizes ((7) below). To formalise this, we fix a surplus function $s(\cdot, \cdot)$ and a distribution of abilities for a countably infinite population, called a "meta-environment." Then, we consider "finite environments" consisting of the first N agents of this population, for various $N < \infty$.

We first formalise, in Lemma 2, the condition on the increasing power effect for coalitions of sufficiently small power and its implications. Then, we state our general result for large but finite populations in Theorem 2.

Lemma 2 *Consider a meta-environment that satisfies*

$$s_{11}(p, q) \ge 0 \quad \forall p \in (0, \bar{p}) \;\; \forall q > 0, \quad \text{for some} \quad \bar{p} > 0. \tag{6}$$

(a) In any stable system of any finite environment, there is at most one coalition whose strength is less than $\bar{p}/2$, which we call a "frivolous coalition" (if it exists). (b) For any integer q, there is $N(q)$ such that in any stable system of any finite environment of size $N > N(q)$, every non-frivolous coalition has at least q members.

Proof. (a) Any two coalitions of power less than $\bar{p}/2$ each, say Z and Z' if they exist, would produce, if merged, a total surplus that exceeds the sum of their respective surpluses, i.e., $S(Z \cup Z') > S(Z) + Z(Z')$ owing to (6) and $s_2(p, q) > 0$ from (1). Thus, it would constitute a profitable deviation for them to merge (with an appropriate imputation rule). Therefore, at most one coalition may have power less than $\bar{p}/2$ in any stable system.

(b) Given a fixed integer q, if there exists a stable system with a non-frivolous coalition whose size is $q' < q$, then the contribution an agent makes by joining this coalition is at least $\underline{c} = \min_{p \in [\bar{p}/2,1]}(s(p, q'+1) - s(p, q')) > 0$ regardless of the agent's ability. Thus, all agents in other coalitions should get a payoff no less than $\underline{c}$ (for otherwise it would constitute a profitable deviation for an agent getting paid less than $\underline{c}$ in some other coalition to join this coalition). This is impossible for large enough N because the possible total surplus in this environment is bounded above by (1). ■

Theorem 2 *Consider two imputation rules* $\mathbf{f}$ *and* $\mathbf{g}$ *in a meta-environment that satisfies (4), (6) and*

$$\lim_{n' \to \infty} \inf_{n > n'} \frac{a_{n+1}}{a_n} > \lim_{q \to \infty} \sup_{p} \frac{s(p, q+2) - s(p, q+1)}{s(p, q+1) - s(p, q)}, \tag{7}$$

where $\mathbf{f}$ *is less equal than* $\mathbf{g}$. *There exists an integer* λ *that satisfies the following property: For any sufficiently small* $\epsilon > 0$ *there is an integer* $N(\epsilon)$ *such that in any finite environment of size* $N > N(\epsilon)$,

(i) any SSS with $\mathbf{f}$ *as its imputation rule consists of at least* λ *coalitions, not counting the coalition with power less than* ϵ *if it exists, and*

(ii) any SSS with $\mathbf{g}$ *as its imputation rule consists of at most* λ *coalitions, not counting the coalition with power less than* ϵ *if it exists.*

Condition (7) means that eventually the coalition size effect on the surplus diminishes more rapidly than the agent's ability. This seems a reasonable assumption, given that in most organisations and activities the returns to size eventually decrease strongly. In this case, the theorem establishes that a less equal stable system consists of at least as many coalitions as a more equal stable system does, as long as the population is large enough, ignoring "frivolous" coalitions of negligible power, which can be interpreted as the residual agents who wish, but are unable, to grow to a larger, proper coalition.

The relationship between intra-coalition inequality and the structure of a stable partition is obtained for quite general environments in this section, and thus is applicable broadly so long as stable systems exist. However, it is difficult to fully characterise the circumstances under which stable systems exist and their precise structure when they do, because these characterisations depend on the details of the environment, such as the ability distribution and the shape of the surplus function. In the next section we address these questions in the class of environments in which the surplus of a coalition increases linearly with its power.

4 When surplus is linear in power

In this section we consider an environment in which the surplus of a coalition $Z \subset \Omega$ is linear in its power conditional on its size. That is,

$$S(Z) \; = \; p(Z) \cdot \zeta(|Z|) \quad \text{where} \quad \zeta'(q) > 0 \quad \text{and} \quad \zeta(|\Omega|) = 1. \tag{8}$$

Note that $\zeta(|\Omega|) = 1$ is a normalisation for expositional convenience.

We start with the case of a countably infinite population. Note that in this case it would constitute a profitable deviation for any coalition of a finite size, say Z_1, to merge with any other coalition, say Z_2, because

$$p(Z_1 \cup Z_2)\zeta(|Z_1 \cup Z_2|) \; > \; p(Z_1)\zeta(|Z_1|) + p(Z_2)\zeta(|Z_2|) \tag{9}$$

as $p(Z_1 \cup Z_2) = p(Z_1) + p(Z_2)$ and $\zeta(|Z_1 \cup Z_2|) \geq \zeta(|Z_i|)$ for $i = 1, 2$, where the inequality is strict for at least one i. In other words, the marginal contribution of each coalition to the total surplus of the merged coalition is larger than its original surplus, strictly for coalitions of any finite size. Therefore, every coalition must be of an infinite size in any stable system. Furthermore, as the marginal contribution of each agent i to any coalition of an infinite size is $a_i \cdot \zeta(\infty) = a_i$, every agent must be paid at least her ability in any stable system. This is because if any agent were paid less than her ability, some other agent should be paid more than her ability in the same coalition and thus, it would be profitable for the coalition to expel the latter agent and share the saved excess payment among the remaining members. By the same token, no coalition may pay any member more than her marginal contribution. Therefore, in any stable system every agent's payoff is equal to her ability, or her maximum possible marginal contribution to a coalition and, conversely, any such system is stable because no profitable deviation exists when everyone is paid their maximum possible marginal contribution, as summarised in the next result.[10]

Proposition 1 *If (8) holds and $|\Omega| = \infty$, then a system is stable if and only if every coalition is of an infinite size and every agent's payoff is equal to her ability.*

In the case of a finite population, on the other hand, (8) implies that (9) holds for any two coalitions in a system and, consequently, that any stable system consists of a grand coalition, Ω. Thus, the grand coalition constitutes a stable system if it pays the agents their marginal contributions, which are equal to their abilities (due to the normalisation that $\zeta(|\Omega|) = 1$), or sufficiently close to those levels so that no group of agents may form a profitable deviation, as stated below.

Proposition 2 *If (8) holds and $|\Omega| < \infty$, then a system is stable if and only if (i) it consists of a grand coalition and (ii) for any subset D of agents $S(D)$ is no more than the sum of their payoffs in the system.*

[10] We use "marginal contribution" of an agent i to mean her contribution to the total surplus as a member of the coalition Z she belongs to, i.e., $a_i \cdot \zeta(|Z|)$ rather than $S(Z) - S(Z \setminus \{i\})$.

The result that stability requires a grand coalition in finite populations stems from the assumption of strictly increasing surplus in coalition size, i.e., $\zeta'(q) > 0$ of (8). If we modify this assumption so that the surplus ceases to increase in coalition size after a certain level,[11] then stable systems for large finite populations are characterised analogously to Proposition 1 above, as stated below.

Proposition 3 *Suppose that (8) holds with the modification that $\zeta'(q) > 0$ for $q < M \in \mathbb{N}$ and $\zeta(q) = 1$ for $q \geq M$. If $|\Omega| > M$ then a system is stable if and only if every coalition is of size M or larger and every agent's payoff is equal to her ability.*

Proof. Consider a system in the environment presumed in the Proposition. If any coalition, say Z, is of a size less than M, then (9) holds for any other coalition Z' in the system and thus, $Z \cup Z'$ would constitute a profitable deviation. Hence, any stable system must consist of coalitions of size M or larger. If any agent, say i, gets a payoff less than a_i in such a system, then she would benefit by joining another coalition, if exists, because her marginal contribution to that coalition is a_i. If there is only one coalition, then there must exist another agent, say j, who gets paid more than her ability and it would be profitable for the coalition to expel j and share the saved excess payment because the total surplus would decrease only by a_j given $|\Omega| > M$. Therefore, every agent i gets a payoff of a_i in any stable system. Conversely, if every agent i gets a payoff of a_i in a system, then no profitable deviation exists because a_i is the maximum possible marginal contribution of agent i to any coalition, and therefore the system is stable. This completes the proof. ■

Fully flexible negotiation of the intra-coalition division of surplus implies that efficiency is achieved in stable systems and that each agent's payoff is equal to her maximum possible marginal contribution. In the environments considered in this section, stable systems always exist and admit a complete characterisation, besides exhibiting the relationship delineated in the previous section between intra-coalition inequality and the number of coalitions in the stable systems.

For the sake of providing a clear illustration, in the remaining part of this section we present a heuristic class of environments where the agent's ability is distributed according to a geometric distribution and coalitions are restricted to choose an imputation rule that can be represented by a single parameter depicting the payoff ratio between any two consecutively ranked members.[12,13]

---

[11] In line with the power in politics interpretation, this case of flat ζ after a threshold size is reminiscent of the party formation problem, where each party will have exactly as many members or candidates as there are posts in a majoritarian government plus patronage system.

[12] Any two systems can be compared unambiguously in terms of internal inequality under this restriction.

[13] Barbera, *et al.* (2014) study stable organisational structures in a related model where each coalition is restricted to choose by voting (hence, dictated by the median ability member) between two imputation rules, namely, the meritocratic rule, according to which the agents get their marginal contribution, and the egalitarian rule, according to which everyone gets the same

4.1 Geometric ability distribution and single parameter imputation rule

Consider a countably infinite population and a geometric distribution of abilities:

$$a_i = (1-a)a^{i-1} \quad \text{for} \quad i \in \mathbb{N}, \quad \text{where} \quad a \in (0,1). \tag{10}$$

Here a reflects the degree of ability differential across agents. Note that multiplication by $(1-a)$ is for normalisation to 1 of the sum of abilities of all agents.

We now capture the vertical inequality of payoffs across ranks with a single parameter ρ, termed the *imputation ratio,* which is the ratio of the surplus share of any coalition member relative to that of the member occupying the rank immediately above.[14] Specifically, each coalition Z is assumed to choose an imputation ratio $\rho \in (0,1)$ as its imputation rule, that is, $f_{r+1}/f_r = \rho$ for all $r = 1, 2, \cdots, |Z|-1$. Note that a lower ρ corresponds to greater inequality.

We retain the assumption that every coalition adopts ability ranking, which is innocuous for our purpose of characterising stable systems because Lemma 1 applies straightforwardly to the current environment as well. Then, denoting the rank of agent i in a coalition Z by $r_i(Z)$, the expected utility of agent i in Z is

$$\begin{aligned} u_i(Z,\rho) &= \frac{S(Z)\cdot \rho^{r_i(Z)-1}}{1+\rho+\cdots+\rho^{|Z|-1}} = \frac{S(Z)(1-\rho)\rho^{r_i(Z)-1}}{1-\rho^{|Z|}} \\ &= S(Z)(1-\rho)\rho^{r_i(Z)-1} \quad \text{if} \quad |Z| = \infty. \end{aligned} \tag{11}$$

Thus, every agent should decide which coalition to join not only on the basis of the coalition's power, $p(Z)$, but also on the basis of her expected rank in the coalition, $r_i(Z)$, and the vertical inequality, ρ.

We now represent a system as $(\pi, \vec{\rho})$ consisting of a partition $\pi = \{Z_1, \cdots, Z_K\}$ of Ω into K coalitions, and a K-vector $\vec{\rho} = (\rho_1, \cdots, \rho_K)$ that specifies one imputation ratio $\rho_k \in (0,1)$ for each coalition $Z_k \in \pi$. Note from (8) that the marginal contribution of an agent i to a coalition's total surplus is $a_i = (1-a)a^{i-1}$ when she joins a coalition of an infinite size. As before, a key observation here is that every agent can obtain this level of payoff. Formally,

share. They also obtain non-segregated groups and the coexistence of different reward norms (meritocratic and egalitarian groups) as inherent features of stable structures. Their modelling of coalitional production is a limit case of (8) where $\zeta(\cdot)$ jumps from 0 to 1 at a certain threshold size. If coalitions were allowed to choose from unrestricted imputation rules, therefore, the logic behind Proposition 3 would imply that, unlike their results, stable structures always exist and are meritocratic. (This is the case in the absence of their tie-breaking rule which stipulates that if an agent gets the same payoff from two coalitions then she prefers the one with a higher average ability. If the tie-breaking rule is imposed on our setting when $\zeta(\cdot)$ jumps from 0 to 1 at a certain threshold size, then a stable system does not exist unless the total number of agents is an integer multiple of the threshold size.)

[14] In the political economy literature it is very common to simplify the distributive views using a single parameter, such as the preferred tax rate in a linear tax system.

Lemma 3 *If an agent i's payoff in a system $(\pi, \vec{\rho})$ is strictly less than $(1-a)a^{i-1}$, then the system is not stable.*

Proof. Let j_1 be the agent whose payoff in the system $(\pi, \vec{\rho})$ is strictly lower than $(1-a)a^{j_1-1}$. Find a sufficiently low $\rho' > 0$ such that $u' = (1-\rho')(1-a)a^{j_1-1}$ exceeds her payoff in $(\pi, \vec{\rho})$. For each $r = 2, 3, \cdots$, one can find an agent, say j_r, whose payoff in the system $(\pi, \vec{\rho})$ falls short of $u' \cdot (\rho')^{r-1}$, maintaining the feature that $j_r < j_{r+1}$. This is because there exists an agent i, where i is arbitrarily large, whose payoff is arbitrarily low in the system either because her rank is arbitrarily low in an infinite coalition, or in the case that there is no coalition of an infinite size, because she is in a coalition of arbitrarily small power. Then, the deviation $D' = \{j_1, j_2, \cdots\}$ with the imputation ratio ρ' is profitable because agent j_r would have a payoff of $(1-\rho')(1-a)(\sum_{n=1}^{\infty} a^{j_n-1})(\rho')^{r-1} > u' \cdot (\rho')^{r-1}$. This proves that $(\pi, \vec{\rho})$ is not stable. ■

Lemma 3 implies that, in any stable system, agent i's payoff is at least $(1-a)a^{i-1}$. In fact, it is *equal* to $(1-a)a^{i-1}$ in any stable system because the maximum possible surplus in the whole economy is $S(\Omega) = (1-a)\sum_{i=1}^{\infty} a^{i-1}$. This further implies that any coalition Z in a stable system must generate a total surplus of $S(Z) = (1-a)\sum_{i \in Z} a^{i-1}$, which is possible only if it is of an infinite size.

The specific partition structure defined below turns out to be crucial. A coalition Z is "K-cyclic" if it consists of every K-th agent starting from a certain agent k, i.e., $Z = \{k, k+K, k+2K, k+3K, \cdots\}$. A "$K$-cyclic partition" is $\pi_K^c = \{Z_1, \cdots, Z_K\}$ where each Z_k is K-cyclic starting from agent k for $k = 1, 2, \cdots, K$.

A "symmetric K-cyclic system" (π_K^c, ρ) where π_K^c is the K-cyclic partition and $\rho = (a^K, \cdots, a^K)$, clearly delivers every agent a payoff that is equal to her ability and thus, constitutes a stable system. In fact, the same conclusion holds so long as each coalition of a system is K-cyclic with an imputation ratio $\rho = a^K$ for some integer K, where the value of K may vary across coalitions. We refer to such a system as a "generalised cyclic" system. For example, in the symmetric 4-cyclic system, if Z_1 and Z_3 merge to form a 2-cyclic coalition and adopt an imputation ratio of a^2, then the new system is a generalised cyclic system.

The next result establishes that a system is stable if and only if it is of this kind. Therefore, organisations with varying norms of internal inequality coexist in a stable system. Moreover, the more unequal internal norms a system displays across the board, the larger is the number of competing organisations that have emerged in the system and the more widely dispersed are the agents' abilities within organisations.

Proposition 4 *A system $(\pi, \vec{\rho})$ is stable if and only if it is a generalised cyclic system. Furthermore, agent i's payoff in any stable system is $(1-a)a^{i-1}$.*

Proof. We have already shown that agent i's payoff is equal to $(1-a)a^{i-1}$ in any stable system. It is clear from (8) and (10) that this is possible for a coalition only if it is K-cyclic and its imputation ratio is a^K for some $K \geq 1$. This proves that if a system is stable then it is a generalised cyclic system. Conversely, any generalised cyclic system is stable because the surplus of any deviation is no more than the sum of the abilities of the deviating members, so no beneficial deviation exists. ■

4.2 Stable partitions under a fixed imputation ratio ρ

In this section we study the stability of partitions in the special case in which ρ is exogenously given (and hence ρ must be respected even in deviations). This environment can be interpreted as a short-run model in which the organisational structure may not be changed quickly due to some institutional reasons.[15] In a political context, for example, even though a party's probability of winning depends on the abilities of the politicians involved, once a party gains power and the various offices have to be filled, at that point the relative payoffs of the various party members depend on their assigned ranks (from president to secretary and so on), and these relative and absolute payoffs take the form of pre-specified rank-dependent remunerations.

The main result of this section is that given $a \in (0, 1)$, if a stable partition exists for any ρ, it is unique and is a K-cyclic partition for some integer K that decreases in ρ, as specified below. Since lower ρ represents higher inequality, this reflects a positive relationship between the vertical inequality and the number of coalitions to be formed endogenously.

Proposition 5 *Suppose that all coalitions adopt a given imputation ratio $\rho \in (0, 1)$. Then, the K-cyclic partition π_K^c is stable if and only if*

$$a^K \leq \rho < \frac{a^{K-1}}{1 + a^{K-1} - a^K}. \tag{12}$$

Furthermore, if a partition is stable for any (a, ρ), it is a K-cyclic partition for some integer K and it is the unique stable partition given (a, ρ). The fraction of ρ values specified by (12) relative to $a^K \leq \rho < a^{K-1}$ approaches 1 as $K \to \infty$.

Proof. See Morelli and Park (2011), an earlier version of this paper. ■

Figure 1 illustrates the areas of the parameter values $(a, \rho) \in (0, 1) \times (0, 1)$ that satisfy (12) and thus, support stable K-cyclic partitions. Between the area for a stable K-cyclic partition and that for a stable $(K-1)$-cyclic partition, there is an area with no stable partition. The fraction of this latter area, relative to the area for a stable K-cyclic partition, converges to 0 as $K \to \infty$.

[15] We share the view that "often, the rewards from joint effort are shared according to rather rigid rules," as conveyed, e.g., by Farrell and Scotchmer (1988) who, focusing on partnerships, analyse the implications of the equal sharing rule on the size of partnerships and welfare.

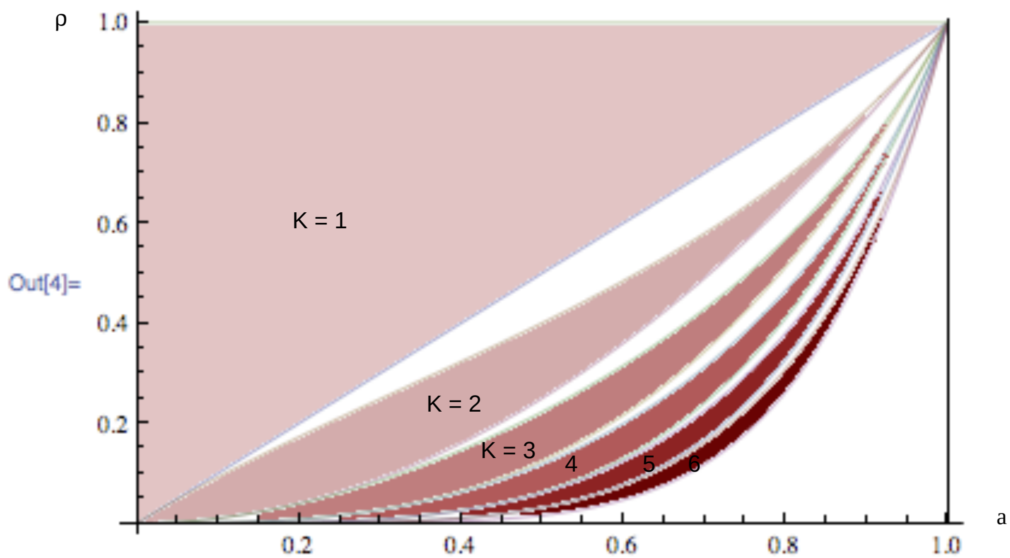

Figure 1

5 Concluding Remarks

In this paper, we have demonstrated some insightful and robust connections between vertical inequality within coalitions and the endogenous formation of coalition structures. In order to best emphasise the connection, we have first established the positive correlation between vertical inequality and the number of organisations for general environments, i.e., for a large class of surplus functions and distributions of ability, and then confirmed the existence of (and provided tight characterisations of) stable outcomes for surplus functions that increase linearly in coalitional power.

Even though our theoretical framework and results are not directly usable for normative analysis, it is important to remark that our results on the relationship between vertical inequality and number of competing organizations can also enter policy or regulation debates. For example, in the recent discussion about the pros and cons of imposing less inequality in pay structures within certain kinds of firms, nobody has mentioned an implication of such a restriction that could emerge from our analysis, namely that such a restriction could lead to greater concentration in the industry.

One limitation of our model is that the value of a coalition does not depend on the partition of the rest of the agents. This limitation may not be critical when coalitions are expected to be of similar strengths such as in our cyclic partitions. In other contexts, however, it may be important for the value of a coalition to reflect the way that other coalitions are composed. In plurality rule elections, for example, it makes a big difference for a coalition expecting 30% of the votes whether the remaining 70% is divided into 7 small parties of 10% each or two parties of 35% each. An extension of the model in which the relative power of any coalition depends not only on the ability of its members but also on some other relevant dimension is in our future research agenda.

We note that our cooperative game theoretic results have the potential to be implemented and extended in a dynamic setting such as the one introduced by Acemoglu, Egorov and Sonin (2008), since the lack of commitment that constitutes their main tenet is also conceptually or implicitly assumed in our core-like cooperative logic. A dynamic stability analysis could therefore constitute the natural next step for this line of research.

In future research it would also be useful to analyse the realistic extension in which abilities are more than one dimensional, to see for example whether stable systems with more groups tend to have a different sorting of ability compositions relative to systems with less competition. In other words, with more than one dimensions of ability and some standard assumption of complementarities or team production, both the distribution of surplus across hierarchical levels and the relationship with equilibrium sorting of groups are important.

A few remarks about the potential empirical relevance of our findings are in order: In the literature on the number of political parties, for example, the leading hypotheses elaborated and tested have all to do with the electoral formula (Duverger's law and Duverger's hypothesis), but it is well documented that even after controlling for the electoral formula the number of parties of different countries varies enormously (think of India and the United States in the set of countries using a majoritarian system and Ireland and Italy within the set of countries using a more proportional system). Within each set of countries with homogeneous electoral institutions, one could in principle verify whether vertical inequality across the major ranks of each party is indeed higher in countries with a larger number of stable parties. Similarly, dividing the US production of goods and services according to several categories, one could evaluate income distribution across ranks in each industry or category and determine if higher hierarchical inequality is correlated with a lower concentration in the sector. However, it is beyond the scope of this paper to determine whether appropriate data exist and to test the prediction while properly controlling other relevant cultural and economic differences across countries/sectors, which we leave for future research.

APPENDIX

We start with two preliminary results. An SSS (π, ρ) is also represented by $(\pi, \mathbf{f})$ where $\mathbf{f} = (f_1, f_2, \cdots)$ is the imputation rule of the largest size coalition in π.

Lemma 4 *In any SSS for an environment with $|\Omega| = \infty$, every coalition has countably infinite members.*

Proof. To reach a contradiction, suppose there is a finite coalition, say Z_1, in an SSS $(\pi, \mathbf{f})$. Note that for any $j \notin Z_1$ we have $s(p(Z_1 \cup \{a_j\}), |Z_1| + 1) - s(p(Z_1), |Z_1|) > s(p(Z_1), |Z_1| + 1) - s(p(Z_1), |Z_1|) > s(p(Z_1), |Z_1|)s_2(p(Z_1), |Z_1| + 1) > 0$ by (1), i.e., $s(p(Z_1 \cup \{a_j\}), |Z_1| + 1) - s(p(Z_1), |Z_1|)$ exceeds a positive number independently of a_j.

If there is an infinite coalition, say $Z_k \in \pi$, since $\lim_{r \to \infty} f_r = 0$, there exists $j \in Z_k$ such that agent j's payoff in the system $(\pi, \mathbf{f})$ is arbitrarily small, in particular, strictly smaller than $\delta = s(p(Z_1 \cup \{a_j\}), |Z_1| + 1) - s(p(Z_1), |Z_1|)$. So, agent j would be better off by joining Z_1 if paid δ and thus, $D = Z_1 \cup \{j\}$ would constitute a profitable deviation because it would not make any member of Z_1 worse off, contradicting the supposed stability of the partition. If all coalitions of π are finite, on the other hand, there are infinitely many coalitions and consequently, there is a coalition of arbitrarily small power, say $Z' \in \pi$, such that the payoff of any agent $j \in Z'$ is smaller than $s(p(Z_1 \cup \{a_j\}), |Z_1| + 1) - s(p(Z_1), |Z_1|)$, leading to an analogous contradiction that $Z_1 \cup \{j\}$ would be a profitable deviation. ■

Lemma 5 *Let $(\{Z_1, Z_2, \cdots, Z_K\}, \mathbf{f})$ be an SSS (where $K = \infty$ is allowed).*

(a) $p(Z_k) > p(Z_{k+1})$ for all $k = 1, 2, \cdots, K$.

(b) If $k < k'$, the member of Z_k at rank r is at least as able as the member of $Z_{k'}$ at the same rank for every $r = 1, 2, \cdots$.

Proof. (a) Suppose to the contrary that $p(Z_{k+1}) \geq p(Z_k)$ for some k. Let i_k and i_{k+1} be the most able member of Z_k an Z_{k+1}, respectively. Then, the deviation $D = \big(Z_{k+1} \backslash \{i_{k+1}\}\big) \cup \{i_k\}$ with the imputation rule $\mathbf{f}$ would be profitable because (i) $i_k < i_{k+1}$ so that $p(D) > p(Z_{k+1}) \geq p(Z_k)$, and (ii) every member in D retains the same rank as in the original system.

(b) Suppose to the contrary that the rank r member of Z_k, say j, is less able than the same rank member of $Z_{k'}$, say i, for some r. Consider the lowest such r. Since $p(Z_k) > p(Z_{k'})$ by part (a), the deviation $D = \big(Z_k \backslash \{j\}\big) \cup \{i\}$ with the imputation rule $\mathbf{f}$ would be profitable because (i) $p(D) > p(Z_k) > p(Z_{k'})$, and (ii) every member in D retains the same rank as in the original system. ■

Proof of Theorem 1. Consider two SSS's, F and G, in an environment that satisfies (4). Assume ability ranking in both F and G by Lemma 1. Recall from Lemma 4 that all coalitions are of an infinite size. Let $\mathbf{f} = (f_1, f_2, \cdots)$ and $\mathbf{g} = (g_1, g_2, \cdots)$ be, respectively, the common imputation rule of F and G.

Suppose, to the contrary, that F is less equal than G, yet F consists of K coalitions denoted by $Y_1, \cdots, Y_K$, while G consists of $L > K$ coalitions denoted by $Z_1, \cdots, Z_L$. For each integer r, let $Y_k|_r$ be the subset of Y_k consisting of all agents ranked r or below in Y_k, for every $k = 1, \cdots, K$; Similarly, let $Z_\ell|_r$ be the subset of Z_ℓ consisting of all agents ranked r or below in Z_ℓ, for every $\ell = 1, \cdots, L$.

First, we establish that

$$\frac{p(Y_k|_r)}{p(Z_\ell|_r)} \text{ is bounded above for every } (k, \ell) \in \{1, \cdots, K\} \times \{1, \cdots, L\}. \tag{13}$$

To show this, observe that stability of F implies for each rank r:

(i) For the subset $Y_k|_r$ to not constitute a profitable deviation together with some imputation rule (e.g., keep the payoff ratio between any two agents in $Y_k|_r$ the same as in F), we need $s(p(Y_k), \infty) \cdot (\sum_{j=0}^{\infty} f_{r+j}) \geq \underline{v} p(Y_k|_r)$ by (1);

(ii) Let $Y_k^1|_r \subset Y_k|_r$ denote the subset consisting of every other member of $Y_k|_r$ starting from the most able member. For the subset $Y_k \setminus Y_k^1|_r$ to not constitute a profitable deviation together with some imputation rule, we need that what $Y_k^1|_r$ collectively gets in Y_k is no more than what it contributes, i.e., $s(p(Y_k), \infty) \cdot (\sum_{j=0}^{\infty} f_{r+2j}) \leq s(p(Y_k), \infty) - s(p(Y_k \setminus Y_k^1|_r), \infty) \leq \bar{v} p(Y_k^1|_r)$ where the last inequality follows because $s_1(\cdot, \cdot)$ is bounded above by $\bar{v}$.

(iii) Similarly, letting $Y_k^2|_r = Y_k|_r \setminus Y_k^1|_r$, for the subset $Y_k \setminus Y_k^2|_r$ to not constitute a profitable deviation together with some imputation rule, we need that what $Y_k|_r^2$ collectively gets in Y_k is no more than what it contributes, i.e., $s(p(Y_k), \infty) \cdot (\sum_{j=0}^{\infty} f_{r+2j+1}) \leq s(p(Y_k), \infty) - s(p(Y_k \setminus Y_k^2|_r), \infty) \leq \bar{v} p(Y_k^2|_r)$.

Combining the inequalities from (ii) and (iii),[16] we get $s(p(Y_k), \infty) \cdot (\sum_{j=0}^{\infty} f_{r+j}) \leq \bar{v} p(Y_k|_r)$. Together with the inequality from (i), we deduce further that

$$\frac{\sum_{j=0}^{\infty} f_{r+j}}{\bar{v}} \leq \frac{p(Y_k|_r)}{s(p(Y_k), \infty)} \leq \frac{\sum_{j=0}^{\infty} f_{r+j}}{\underline{v}}. \tag{14}$$

Analogous arguments for G establishes

$$\frac{\sum_{j=0}^{\infty} g_{r+j}}{\bar{v}} \leq \frac{p(Z_\ell|_r)}{s(p(Z_\ell), \infty)} \leq \frac{\sum_{j=0}^{\infty} g_{r+j}}{\underline{v}}. \tag{15}$$

Since (3) implies that $(\sum_{j=0}^{\infty} f_{r+j})/(\sum_{j=0}^{\infty} g_{r+j}) < 1$ for sufficiently large r, (14) and (15) establish (13).

[16] Note that we do not obtain the same condition by by requiring $Y_k \setminus Y_k|_r$ to not constitute a profitable deviation, because $|Y_k \setminus Y_k|_r| < \infty$.

Now, for each r, let $i_k(r) \in \Omega$ be the rank r agent in Y_k, and define $\underline{i}(r) = \min\{i_1(r), i_2(r), \cdots, i_K(r)\}$ and $\bar{i}(r) = \max\{i_1(r), i_2(r), \cdots, i_K(r)\}$. Similarly, letting $j_\ell(r)$ be the rank r agent in Z_ℓ, define $\underline{j}(r) = \min\{j_1(r), j_2(r), \cdots, j_L(r)\}$ and $\bar{j}(r) = \max\{j_1(r), j_2(r), \cdots, j_L(r)\}$. Then, Lemma 5 (b) implies that

$$\underline{i}(r) < rK \leq \bar{i}(r) \quad \text{and} \quad \underline{j}(r) < rL \leq \bar{j}(r),$$

which in turn imply that $\bar{j}(r) - \underline{i}(r) > (L-K)r$ so that $\bar{j}(r) - \underline{i}(r) \to \infty$ as $r \to \infty$.

By (4), one can find α such that $\limsup \frac{a_{n+1}}{a_n} < \alpha < 1$ and, consequently, $\frac{a_{n+1}}{a_n} < \alpha$ for all sufficiently large n. Then, for sufficiently large r, we would have $\sum_{i \in Y_k|_r} a_i > a_{\underline{i}(r)}$ and $\sum_{i \in Z_\ell|_r} a_i < a_{\bar{j}(r)}/(1-\alpha)$ for $k = 1$ and $\ell = L$ by Lemma 5 (b), contradicting (13) because $a_{\bar{j}(r)} < a_{\underline{i}(r)} \alpha^{\bar{j}(r) - \underline{i}(r)}$ so that $p(Y_k|_r)/p(Z_\ell|_r) > (1-\alpha) a_{\underline{i}(r)}/a_{\bar{j}(r)} > (1-\alpha)/\alpha^{\bar{j}(r) - \underline{i}(r)} \to \infty$ as $r \to \infty$. ■

Proof of Theorem 2. We assume $\epsilon < \bar{p}/2$ below where $\bar{p}$ is as in (6). Then, by Lemma 2 (a), each SSS may have at most one frivolous coalition and consequently, there may be no more than $\kappa = 2/\bar{p} + 1$ coalition in any SSS.

Fix a pair of integers K and L such that $0 \leq K - 1 < L \leq \kappa$. For each $N \in \mathbb{N}$ find all SSS's (if exist), $F^N = (\pi^N_{\mathbf{f}}, \mathbf{f})$ and $G^N = (\pi^N_{\mathbf{g}}, \mathbf{g})$ where $\pi^N_{\mathbf{f}}$ consists of K coalitions denoted by $Y^N_1, \cdots, Y^N_K$, while $\pi^N_{\mathbf{g}}$ consists of L coalitions denoted by $Z^N_1, \cdots, Z^N_L$, including the frivolous ones. If, for any such pair of K and L, there does not exist any SSS with **f** or **g** as its imputation rule for all sufficiently large N, then the Theorem holds trivially. Hence, assume otherwise in the remainder of this proof.

The key step of the proof is the following result:

[*] There do not exist two sequences $N_n \to \infty$ and $N'_n \to \infty$ for which one can find two sequences of SSS's, $(F^{N_n})_n$ and $(G^{N'_n})_n$, with the following property: The limits $\lim_{n\to\infty} p(Y^{N_n}_k)$ and $\lim_{n\to\infty} p(Z^{N'_n}_\ell)$ exist for all $k = 1, 2, \cdots, K$ and $\ell = 1, 2, \cdots, L$, and furthermore,

$$K^* = \#\{k | \lim_{n\to\infty} p(Y^{N_n}_k) > 0\} \ < \ L^* = \#\{\ell | \lim_{n\to\infty} p(Z^{N'_n}_\ell) > 0\}. \tag{16}$$

To prove this, suppose to the contrary that two sequences $(F^{N_n})_n$ and $(G^{N'_n})_n$ exist with the above property. By Lemma 2, $K^* \in \{K-1, K\}$ and $L^* \in \{L-1, L\}$. By convention of indexing coalitions, $\lim_{n\to\infty} p(Y^{N_n}_k) > 0$ precisely for $k = 1, \cdots, K^*$, and $\lim_{n\to\infty} p(Z^{N'_n}_\ell) > 0$ precisely for $\ell = 1, \cdots, L^*$.

If $K^* = K - 1$, for any integer r, there is $n(r)$ large enough so that the most able member of the weakest coalition $Y^{N_n}_K$ is weaker than $a_{K^* r}$ for all $n \geq n(r)$, and thus,

$$\underline{i}(r) = \min\{i | i \text{ is rank } r \text{ agent in some coalition in } F^{N_n}\} \leq K^* r$$

for all $n \geq n(r)$. This inequality holds straightforwardly when $K^* = K$ as well. In addition, by an analogous reasoning,

$$\bar{j}(r) = \max\{j | j \text{ is rank } r \text{ agent in } Z_\ell^{N'_n} \text{ for some } \ell \leq L^*\} \geq L^* r$$

for all N'_n. Since there is α such that $\limsup \frac{a_{n+1}}{a_n} < \alpha < 1$ for all large n by (4), we can find subsequences $(F^{\hat{N}_r})$ and $(G^{\hat{N}'_r})_r$ such that $p(Y_k^{\hat{N}_r}|_r) > a_{\underline{i}(r)}$ with $\underline{i}(r) \leq K^* r$ for some $Y_k^{\hat{N}_r}$, $k \leq K^*$, and $p(Z_\ell^{\hat{N}'_r}|_r) < a_{\bar{j}(r)}/(1-\alpha)$ with $\bar{j}(r) \geq L^* r$ for some $Z_\ell^{\hat{N}'_r}$, $\ell \leq L^*$, where $Z|_r$ denotes the the subset of a coalition Z consisting of all members of Z ranked r or below. Note that $k = 1$ and $\ell = L^*$ by Lemma 5. Since, $a_{\bar{j}(r)} < a_{\underline{i}(r)} \alpha^{\bar{j}(r) - \underline{i}(r)}$ and $\bar{j}(r) - \underline{i}(r) \geq (L^* - K^*) r \to \infty$ as $r \to \infty$ by (16), it further follows that $p(Y_1^{\hat{N}_r}|_r)/p(Z_{L^*}^{\hat{N}'_r}|_r) > (1-\alpha)/\alpha^{\bar{j}(r)-\underline{i}(r)} \to \infty$ as $r \to \infty$.

Next, for $Y_k^{\hat{N}_r}|_r$ not to constitute a beneficial deviation in $F^{\hat{N}_r}$, we need

$$s(p(Y_k^{\hat{N}_r}), |Y_k^{\hat{N}_r}|) \cdot (\sum_{j=r}^{|Y_k^{\hat{N}_r}|} f_j) \Big/ (\sum_{j=1}^{|Y_k^{\hat{N}_r}|} f_j) \geq \underline{v} p(Y_k^{\hat{N}_r}|_r) \quad \forall r. \tag{17}$$

For each $i \in Z_\ell^{\hat{N}'_r}|_r$, consider the potential deviation $Z_\ell^{\hat{N}'_r} \setminus \{i\}$. For this not to be beneficial, we need $s(p(Z_\ell^{\hat{N}'_r}), |Z_\ell^{\hat{N}'_r}|) \cdot g_i / (\sum_{j=1}^{|Z_\ell^{\hat{N}'_r}|} g_j) \leq s(p(Z_\ell^{\hat{N}'_r}), |Z_\ell^{\hat{N}'_r}|) - s(p(Z_\ell^{\hat{N}'_r} \setminus \{i\}), |Z_\ell^{\hat{N}'_r}| - 1) = s(p(Z_\ell^{\hat{N}'_r}), |Z_\ell^{\hat{N}'_r}|) - s(p(Z_\ell^{\hat{N}'_r} \setminus \{i\}), |Z_\ell^{\hat{N}'_r}|) + s(p(Z_\ell^{\hat{N}'_r} \setminus \{i\}), |Z_\ell^{\hat{N}'_r}|) - s(p(Z_\ell^{\hat{N}'_r} \setminus \{i\}), |Z_\ell^{\hat{N}'_r}| - 1) < \bar{v} a_i + s(p(Z_\ell^{\hat{N}'_r} \setminus \{i\}), |Z_\ell^{\hat{N}'_r}|) - s(p(Z_\ell^{\hat{N}'_r} \setminus \{i\}), |Z_\ell^{\hat{N}'_r}| - 1)$. Summing the final inequality over all $i \in Z_\ell^{\hat{N}'_r}|_r$, we get

$$\begin{aligned} & s(p(Z_\ell^{\hat{N}'_r}), |Z_\ell^{\hat{N}'_r}|) \cdot (\sum_{j=r}^{|Z_\ell^{\hat{N}'_r}|} g_j) \Big/ (\sum_{j=1}^{|Z_\ell^{\hat{N}'_r}|} g_j) \\ & \quad < \bar{v} p(Z_\ell^{\hat{N}'_r}|_r) + \sum_{i \in Z_\ell^{\hat{N}'_r}|_r} \big(s(p(Z_\ell^{\hat{N}'_r} \setminus \{i\}), |Z_\ell^{\hat{N}'_r}|) - s(p(Z_\ell^{\hat{N}'_r} \setminus \{i\}), |Z_\ell^{\hat{N}'_r}| - 1) \big). \end{aligned} \tag{18}$$

At this point, observe that $|Z_\ell^{\hat{N}'_r}| \to \infty$ as $r \to \infty$, because otherwise, i.e., if $|Z_\ell^{\hat{N}'_r}|$ is bounded above as $r \to \infty$, then the marginal contribution of joining the coalition $Z_\ell^{\hat{N}'_r}$ would be bounded away from 0 while the lowest payoff among all agents vanishes as $r \to \infty$ so that it would be a beneficial deviation for the lowest paid agent to join $Z_\ell^{\hat{N}'_r}$. By taking a subsequence if necessary, we may assume that $|Z_\ell^{\hat{N}'_r}| > r(M+1)$ for any finite integer M. In particular, denoting supremum of the fraction on the RHS of (7) by ξ_q for each q, set M to be an integer such that the product of any consecutive $M-2$ values, $\xi_{j+1}\cdots\xi_{j+M-2}$, is less than $\alpha^L/2$ for sufficiently large j. Then, we have

$$M_r = \max\{n\in\mathbb{N} | n \le |Z_\ell^{\hat{N}'_r}|/r\} > M \quad \text{and} \quad h_r = |Z_\ell^{\hat{N}'_r}| - rM_r < r. \tag{19}$$

Now, we prove that the distance between $\bar{j}(r)$ and the most able rank r agent in G^N, denoted by $\underline{j}(r)$, is bounded above. To see this, first observe that $\underline{j}(r) \in Z_1^{\hat{N}'_r}$ and $\bar{j}(r) \in Z_{L^*}^{\hat{N}'_r}$ by Lemma 5. For $\{\underline{j}(r)\}$ not to be a beneficial deviation, the payoff to $\underline{j}(r)$ in $Z_1^{\hat{N}'_r}$ should be no less than $\underline{v}p(\{a_{\underline{j}(r)}\})$. For $Z_{L^*}^{\hat{N}'_r} \setminus \{\bar{j}(r)\}$ not to be a beneficial deviation, the payoff to $\bar{j}(r)$ in $Z_{L^*}^{\hat{N}'_r}$ should be no more than $\bar{v}p\{a_{\bar{j}(r)}\})$ for large enough r. Therefore, if $\bar{j}(r) - \underline{j}(r) \to \infty$ as $r\to\infty$, as it would follow that $a_{\underline{j}(r)}/a_{\bar{j}(r)} \to \infty$ as $r \to \infty$, we would have established that the payoff ratio between $\underline{j}(r)$ and $\bar{j}(r)$ must explode as $r\to\infty$. But, this is impossible as the following calculation shows:

$$\frac{g_r \cdot s(p(Z_1^{\hat{N}'_r}), |Z_1^{\hat{N}'_r}|)}{(\sum_{j=1}^{|Z_1^{\hat{N}'_r}|} g_j)} \Bigg/ \frac{g_r \cdot s(p(Z_{L^*}^{\hat{N}'_r}), |Z_{L^*}^{\hat{N}'_r}|)}{(\sum_{j=1}^{|Z_{L^*}^{\hat{N}'_r}|} g_j)} \to \frac{s(p(Z_1^{\hat{N}'_r}), \infty)}{s(p(Z_{L^*}^{\hat{N}'_r}), \infty)} \in (0,\infty) \text{ as } r\to\infty.$$

This establishes that $\bar{j}(r) - \underline{j}(r)$ is bounded above, say by B.

Then, because $\underline{j}(r) \le Lr$, it follows that $p(Z_{L^*}^{\hat{N}'_r}|_r) > a_{Lr+B}$ and consequently, $p(Z_{L^*}^{\hat{N}'_r}|_r) \ge \phi\alpha^{Lr}$ for all large enough r for some $\phi > 0$ because $\liminf \frac{a_{n+1}}{a_n} > 0$ by (7). Hence,

$$\begin{aligned}
&\sum_{i\in Z_{L^*}^{\hat{N}'_r}|_r} \frac{s(p(Z_{L^*}^{\hat{N}'_r}\setminus\{i\}), |Z_{L^*}^{\hat{N}'_r}|) - s(p(Z_{L^*}^{\hat{N}'_r}\setminus\{i\}), |Z_{L^*}^{\hat{N}'_r}|-1)}{p(Z_{L^*}^{\hat{N}'_r}|_r)} \\
\le\ & \frac{|Z_{L^*}^{\hat{N}'_r}|\cdot\xi_1\cdots\xi_{|Z_{L^*}^{\hat{N}'_r}|-2}}{\phi\alpha^{Lr}} \times \sup_p \big(s(p,2)-s(p,1)\big) \\
=\ & \frac{(h_r+M_r)\xi_1\xi_2\cdots\xi_{h_r+M_r}}{\phi\alpha^L} \times \frac{(h_r+2M_r)\xi_{h_r+M_r+1}\cdots\xi_{h_r+2M_r}}{(h_r+M_r)\alpha^L} \times\cdots \\
& \times\frac{(h_r+rM_r)\xi_{h_r+(r-1)M_r+1}\cdots\xi_{h_r+rM_r-2}}{(h_r+(r-1)M_r)\alpha^L} \times \sup_p \big(s(p,2)-s(p,1)\big) \\
\to\ & 0 \quad \text{as} \quad r\to\infty.
\end{aligned}$$

The convergence ensues because i) $(h_r+M_r)\xi_1\xi_2\cdots\xi_{h_r+M_r}$ is bounded,[17] ii) $\frac{h_r+r'M_r}{h_r+(r'-1)M_r}$ is bounded above by 2 for $r'\ge 2$, and iii) $\frac{\xi_{h_r+(r'-1)M_r+1}\cdots\xi_{h_r+r'M_r-2}}{\alpha^L} < \frac{1}{2}$ for all large

[17]This is because for any $\beta\in(0,1)$, we have $n\beta^n = \beta(2\beta)(3\beta/2)\cdots(n\beta/(n-1)) \to 0$ as $n\to\infty$ as $n/(n-1)\to 1$.

r' by definition of M_r. and iv) $\sup_p \big(s(p,2) - s(p,1)\big)$ is bounded above by (1). Therefore, from (18) we deduce that $s(p(Z_{L^*}^{\hat{N}_r'}), |Z_{L^*}^{\hat{N}_r'}|) \cdot (\sum_{j=r}^{|Z_{L^*}^{\hat{N}_r'}|} g_j)/(\sum_{j=1}^{|Z_{L^*}^{\hat{N}_r'}|} g_j) < (\bar{v}+\eta)p(Z_{L^*}^{\hat{N}_r'}|_r)$ for some $\eta > 0$.

Together with (17), we further deduce that

$$\frac{p(Y_1^{\hat{N}_r}|_r)}{p(Z_{L^*}^{\hat{N}_r'}|_r)} < \frac{(\bar{v}+\eta)\cdot s(p(Y_1^{\hat{N}_r}), |Y_1^{\hat{N}_r}|) \cdot (\sum_{j=r}^{|Y_1^{\hat{N}_r}|} f_j) \cdot (\sum_{j=1}^{|Z_{L^*}^{\hat{N}_r'}|} g_j)}{\underline{v} \cdot s(p(Z_{L^*}^{\hat{N}_r'}), |Z_{L^*}^{\hat{N}_r'}|) \cdot (\sum_{j=1}^{|Y_1^{\hat{N}_r}|} f_j) \cdot (\sum_{j=r}^{|Z_{L^*}^{\hat{N}_r'}|} g_j)}.$$

The RHS of this inequality is bounded because $(\sum_{j=r}^{|Y_1^{\hat{N}_r}|} f_j)/(\sum_{j=r}^{|Z_{L^*}^{\hat{N}_r'}|} g_j)$ is bounded above contradicting the earlier finding that $p(Y_1^{\hat{N}_r}|_r)/p(Z_{L^*}^{\hat{N}_r'}|_r) \to \infty$. This proves [*].

Finally, to prove the Theorem, suppose to the contrary that there does not exist an integer λ that satisfies the properties (i) and (ii) stated in the Theorem. Note that the property (i) holds trivially when $\lambda = 1$; and that if (i) holds for some $\lambda > 2$ then so it does for $\lambda - 1$. Hence, let $\bar{\lambda}$ be the maximum value of λ for which (i) holds, which must exist due to Lemma 2. This would mean that for any sufficiently small ϵ, an SSS can be found for arbitrarily large N such that its imputation component is $\mathbf{f}$ and it consists of exactly $\bar{\lambda}$ coalitions, not counting the coalition with power less than ϵ.

Then, by supposition that the Theorem is false, property (ii) does not hold for $\bar{\lambda}$, i.e., for arbitrarily small ϵ, an SSS can be found for arbitrarily large N such that its imputation component is $\mathbf{g}$ yet it consists of strictly more than $\bar{\lambda}$ coalitions, not counting the coalition with power less than ϵ. Consider such a sequence $\epsilon_n \to 0$ as $n \to \infty$ and for each ϵ_n find such an SSS $G^{N_n'}$ with a frivolous coalition whenever possible. Note that if $G^{N_n'}$ is without a frivolous coalition for some n, then so is $G^{N_n'}$ for all larger n.

Similarly, for each ϵ_n find an SSS F^{N_n} such that its imputation component is $\mathbf{f}$ and it consists of exactly $\bar{\lambda}$ coalitions, not counting the coalition with power less than ϵ_n. Again, find such an SSS F^{N_n} with a frivolous coalition whenever possible. Note that if F^{N_n} is without a frivolous coalition for some n, then so is F^{N_n} for all larger n. This would mean that we can find two sequences of SSS's $(F^{N_n})_n$ and $(G^{N_n'})_n$ that satisfies the property stated in [*], with $K^* = \bar{\lambda}$, a contradiction. ■

References

[1] Acemoglu D., G. Egorov, and K. Sonin (2008): "Dynamics and Stability of Constitutions, Coalitions and Clubs," *NBER* Working Paper 14239.

[2] Akerlof (1997): "Social Distance and Social Decision," *Econometrica,* 65, 1005-1028.

[3] Aumann, R.J., and J.H. Drèze (1974): "Cooperative Games with Coalition Structures", *International Journal of Game Theory,* 3, 217-37.

[4] Bandyopadhyay, S., K. Chatterjee, and T. Sjöström (2009): "Pre-electoral Coalitions and Post-election Bargaining," Mimeo.

[5] Banerjee, S., H. Konishi and T. Sömnez (2001): "Core in a Simple Coalition Formation Game," *Social Choice and Welfare,* 18, 135-53.

[6] Barbera, S., C. Bevia and C. Ponsati (2014): "Meritocracy, Egalitarianism and the Stability of Majoritarian Organizations," Barcelona GSE, Working Paper #737;

[7] Bogomolnaia A, and M.O. Jackson (2002): "The Stability of Hedonic Coalition Structures," *Games and Economic Behavior*, 38, 201-30.

[8] Casella, A. (2001): "The Role of Market Size in the Formation of Jurisdictions", *Review of Economic Studies,* 68, 83-108.

[9] Casella, A. and J. Feinstein (2002): "Public Goods in Trade: On the Formation of Markets and Jurisdictions", *International Economic Review*, 43, 437-62.

[10] Damiano, E., Hao Li, and W. Suen (2010): "First in Village or Second in Rome?" *International Economic Review,* 51, 263-88.

[11] Dhillon, Amrita (2005): "Political parties and coalition formation." In: G Demange, G and M.Wooders, M. (Eds), *Group formation in economics: Networks, Clubs and Coalitions*, Cambridge University Press, New York, 289-311.

[12] G. Demange (1994): "Intermediate preferences and stable coalition structures," *Journal of Mathematical Economics,* 23, 45-58.

[13] Ellickson, B. (1973): "A Generalization of the Pure Theory of Public Goods," *American Economic Review*, 63, 417-32.

[14] Ellickson, B., B. Grodal, S. Scotchmer and W.R. Zame (1999): "Clubs and the Market," *Econometrica*, 67, 1185-217.

[15] Farrell, J., and S. Scotchmer (1988): "Partnerships", *Quarterly Journal of Economics,* 10, 279-97.

[16] Garicano, L., and E. Rossi-Hansberg (2006): "Organization and Inequality in a Knowledge Economy," *Quarterly Journal of Economics,* 121, 1383-1435.

[17] Greenberg, J., and S. Weber (1986): "Strong Tiebout Equilibrium under Restricted Preferences Domain," *Journal of Economic Theory*, 38, 101-117.

[18] Guesnerie, R., and C. Oddou (1981): "Second best taxation as a game," *Journal of Economic Theory,* 25, 67-91.

[19] Herrera, H., and M. Morelli (2010): "Turnout and Power Sharing," mimeo.

[20] Jehiel, P., and S. Scotchmer (2001), "Constitutional Rules of Exclusion in Jurisdiction Formation", *Review of Economic Studies* 68, 393-413.

[21] Konishi, H., Le Breton, M., and S. Weber (1998): "Equilibrium in a Finite Local Public Goods Economy," *Journal of Economic Theory,* 79, 224-244.

[22] Le Breton, M., Ortuno-Ortin and S. Weber (2008): "Gamson's Law and Hedonic Games," *Social Choice and Welfare,* 30, 57-67.

[23] Legros, P., and A. Newman (1996): "Wealth Effects, Distribution, and the Theory of Organization," *Journal of Economic Theory,* 70, 312-41.

[24] Levy, G. (2004): "A model of political parties," *Journal of Economic theory*, 115, 250-77.

[25] Mattozzi, A., and A. Merlo (2010): "Mediocracy", mimeo, University of Pennsylvania and Universitat Autonoma de Barcelona.

[26] Milchtaich, I., and E. Winter (2002): "Stability and Segregation in Group Formation," *Games and Economic Behavior,* 38, 318-46.

[27] Morelli, M. (2004): "Party Formation and Policy Outcomes under Different Electoral Systems," *Review of Economic Studies,* 71, 829-853.

[28] Morelli, M., and I.-U. Park (2011): "Internal Hierarchy and Stable Coalition Structures," Economics Working Papers ECO2011/20, European University Institute.

[29] Persico, N., J.C. Rodriguez and D. Silverman (2008): "Factions and Political Competition", mimeo, New York University.

[30] Piccione, M., and R. Razin (2011): "Coalition formation under power relations," *Theoretical Economics*

[31] Scarf, H.E. (1967): “The Core of N Person Game,” *Econometrica*, 35, 50-69.

[32] Tiebout, C.M. (1956): “A Pure Theory of Local Expenditures,” *Journal of Political Economy*, 64, 416-24.

[33] Watts, A. 2007.: “Formation of Segregated and Integrated Groups,” *International Journal of Game Theory,* 35, 505-19.

[34] Westhoff, F. (1977): “Existence of Equilibria in Economies with a Local Public Good,” *Journal of Economic Theory,* 14, 84-112.

[35] Willis, R., and S. Rosen (1979): “Education and Self-Selection,” *Journal of Political Economy,* 87, S7-36.

[36] Wooders, M. (1978): “Equilibria, the Core, and Jurisdiction Structures in Economies with a Local Public Good,” *Journal of Economic Theory,* 18, 328-48.

[37] Yi, S. (1996): “Endogenous formation of customs unions under imperfect competition: open regionalism is good” *Journal of International Economics,* 41, 153-177.

[38] Zame, W.R. (2007): “Incentives, Contracts, and the Markets: a General Equilibrium Theory of Firms,” *Econometrica*, 75, 1453-500.