EPRU Working Paper Series 2009-07

Economic Policy Research Unit
Department of Economics
University of Copenhagen
Øster Farimagsgade 5, Building 26
DK-1353 Copenhagen K
DENMARK
Tel: (+45) 3532 4411
Fax: (+45) 3532 4444
Web: http://www.econ.ku.dk/epru/

Club-in-the-Club: Reform under Unanimity

Erik Berglöf, Mike Burkart,
Guido Friebel, and Elena Paltseva

T

ISSN 0908-7745

Club-in-the-Club: Reform under Unanimity*

Erik Berglöf†, Mike Burkart‡, Guido Friebel§, and Elena Paltseva¶

July 2009

Abstract

In many organizations, decisions are taken by unanimity giving each member veto power. We analyze a model of an organization in which members with heterogenous productivity privately contribute to a common good. Under unanimity, the least efficient member imposes her preferred effort choice on the entire organization. In the presence of externalities and an incomplete charter, the threat of forming an "inner organization" can undermine the veto power of the less efficient members and coerce them to exert more effort. We also identify the conditions under which the threat of forming an inner organization is executed. Finally, we show that majority rules effectively prevent the emergence of inner organizations.

Key words: organizations, club good, voting rules, EU integration
JEL codes: D2, D7, P4

*We have benefited from comments by Tore Ellingsen, Chloe Le Coq, Sergei Guriev and seminar participants in Berkeley, Moscow (CEFIR), Munich, Oxford, Stockholm (SITE) and Toulouse(IDEI). Financial support from the Bank of Sweden Tercentenary Foundation (Burkart) and Jan Wallander's and Tom Hedelius Research Foundation (Paltseva) is gratefully acknowledged. All remaining errors are our own.

†EBRD, ECGI and CEPR

‡Stockholm School of Economics, London School of Economics, CEPR, ECGI and FMG

§Goethe-Universität Frankfurt, IDEI Toulouse and CEPR

¶Copenhagen University, corresponding author: elena.paltseva@econ.ku.dk

1 Introduction

Coordination of individual actions is the core problem that any society must solve to assure the well-being of its members. The greater part of the economics literature focuses on markets; arguably, organizations are an equally important coordination mechanism. When studying organizations, economists typically presuppose the existence of a governance system consisting of rules, penalties or transfers. Club theory, for instance, assumes that there is a system of transfers, taxes and entry fees that can be used to make members of a club behave in accordance with the common interest (e.g., Cornes and Sandler, 1996). Similarly, organization theory presupposes the existence of a principal who coordinates the members of an organization through the use of various monetary and non-monetary instruments (e.g., Milgrom and Roberts, 1992).

In many circumstances, cooperation and organizations exist even if there is no comprehensive governance system. Agents who share a common goal can form a "loosely knit" group. For example, sovereign states may come together to coordinate their actions in specific areas, such as economic policies, protection of the environment, or defence. In such situations, there is *a priori* no structure in place that determines how decisions are taken: the member states must first sort out how to decide. There is initially no other decision rule but unanimity, as pointed out by Rousseau.[1]

Unanimity grants each member of an organization a veto right, thereby protecting her against coercion or what de Tocqueville (1835) called the "tyranny of the majority". But the flipside of unanimity is slow and inflexible decision-making and underprovision of the common good. Heterogeneity is key here: Members who are less committed or less productive can veto any proposal to increase contributions ("effort") to the common good. The problem of holding back other, more productive, members becomes particularly severe when there are complementarities between the members' contributions. In the presence of such "weakest-link" effects, a member who invests too little, limits the amount of the common good for the entire organization.

We argue that organizations operating under the unanimity rule can, nonetheless, provide more common goods than what their least committed members would prefer. The mechanism that can overcome the veto power of the least committed members is the threat of forming a "club-in-the-club". To develop our argument we analyze the

[1] "Indeed, if there were no prior convention, where, unless the election were unanimous, would be the obligation on the minority to submit to the choice of the majority? How have a hundred men who wish for a master the right to vote on behalf of ten who do not? The law of majority voting is itself something established by convention, and presupposes unanimity, on one occasion at least" (Rousseau, 1762).

provision of a common good by an organization in which decisions are taken by unanimity. The club good is produced through a Leontief technology with each member's effort as the inputs; effort should be broadly interpreted as any costly contribution to a common good. The members differ in terms of their effort cost. We exclude transfer mechanisms that can induce club members to exert more effort than they individually prefer. Given each member has a veto power, one would expect that the common good provision is determined by the "weakest link", i.e., the member with the highest cost of effort.

In this setting, we show that the mere possibility of forming a club-in-the-club, to which only high effort providers have access, can increase the amount of the good provided by the entire organization. When staying outside is costly, weaker members increase their effort in order to avoid that the inner organization forms. The more committed members may then prefer not to execute the threat: they may be better off having single membership in the initial club at an increased effort, compared to having dual membership in both the initial and the inner club. Thus, the threat of forming a club-in-the-club limits the leverage less committed members have by virtue of their veto power, and unanimity does not necessarily lead to stagnation.

Alternative reasons why less committed members may refrain from executing their veto rights such as reputation or log-rolling rely on repeated interaction and low time-discount rates. By contrast, our mechanism functions in a static model. The threat of forming a club-in-the-club has, however, bite only if two conditions hold. First, the formation of an inner club is not subject to the organization's decision rule, because otherwise less productive members could simply veto it. Hence, the very constitutional incompleteness that makes it hard to incentivize and coordinate members, can help to reform the club through the threat of an inner club. Second, the club-in-the-club constitutes a threat only if either the value from belonging to the initial club decreases when a club-in-the-club actually forms, or if inner club membership is per se valuable.

The threat of an inner club can increase organization-wide effort, but it may also be executed in equilibrium. We show that an inner club can only exist when the original club is sufficiently heterogeneous, and when the deadweight loss of forming an inner club is sufficiently small. Hence, the possibility of clubs-in-clubs may lead to *more* integration, if the off-equilibrium threat makes all members increase their effort, but it can also lead to less integration, if the threat is executed in equilibrium.

The emergence of an inner club is determined by the trade-off between two potential costs. On the one hand, an inner club reduces the benefits of the initial organization. On the other hand, in the absence of an inner organization more efficient members are held back in their effort choice. The latter effect is weaker in organizations operating

under majority rule as the decisive member is more productive than under unanimity. This suggests that a majority rule could be a remedy against disintegration. We show that this is indeed the case, and that supermajority rules often suffice to prevent inner clubs. The required majority threshold depends on the characteristics of the organization and is lower when members are more heterogenous.

The logic of our theory applies to organizations that do not (yet) have a governance structure in place that resolves or alleviates incentive problems through transfers and penalties. Such organizations can be found in various spheres. Faculty members in a university department may form a research group to have a platform to discuss ideas; home owners may form a residential community association to improve the local public goods provision, such as safety and leisure facilities; small groups of concerned citizens may decide to form an NGO to protect the environment or fight racism; national football leagues may want to organize a European-wide tournament. The challenge for these loosely knit groups and organizations is to find instruments to incentivize the members to contribute to the common goal. On the most aggregate level, member countries may want to further advance European integration, which started with the idea of coordinating steal and coal production and then expanded to include many other economic as well as political areas. Indeed, several important episodes of the European integration experience seems to be well in line with the main ideas of our theory. We discuss these parallels at the end of the paper.

The threat of an inner organization and its possible execution parallels arguments put forward in the literature on secession. Our mechanism is related to the one in Buchanan and Faith (1987) on "internal exit" as an alternative to "voting with one's feet". In their theory, the optimal tax rate is derived as the one that maximizes revenues without triggering secession. Our results are complementary to Buchanan and Faith, albeit derived in a different framework (privately costly efforts rather than taxes). We characterize all possible outcomes associated with an internal exit threat (or club-in-the-club) and show that the threat may or may not be executed. Sufficiently high heterogeneity of members is a necessary condition for the threat to be executed and reforms of the voting scheme help to overcome the risk of internal exits. In Bolton and Roland (1997) secession involves lost economies of scale in public good provision, but avoids the "tyranny of the majority" by creating more homogenous political entities.[2] Our focus is, however, not so much on secessions – that is, the complete separation of federations – but rather on the creation of costly internal struc-

[2]Bordignon and Brusco (2001) point out that constitutionally defined secession rights involve a trade-off; they reduce the cost of an actual break-up ex post, but they increase the likelihood of break-up.

tures. Further, we emphasize the potential function of the inner organization threat as a mechanism to discipline less committed members. This contrasts with Gradstein (2004) who argues that secession rights, while protecting minority rights, involve inefficiencies in bargaining processes. In our model, internal threats can increase efficiency (because they can induce higher effort), or decrease it (as the formation of an inner club entails a deadweight loss). An additional distinction is that the above papers consider majority voting, while our main argument concentrates on the unanimity rule – the natural rule for organizations with highly incomplete constitutions.

Our model indicates that organizations may choose to abandon unanimity and subject their members to the will of the majority. This result is related to a growing literature analyzing how constitutions form, in particular, what determines the voting rules of a society. Aghion and Bolton (2003) identify a trade-off between minority protection and flexibility. To adapt to changes, a society must offer transfers to some individuals to prevent them from exercising their veto right. Hence, a society may under the veil of ignorance decide to replace unanimity by some type of majority voting. Messner and Polborn (2004) take a complementary view and show why societies may opt for supermajorities rather than simple majority voting. In their model, young people, who vote today over tomorrow's decision rule, anticipate that they will benefit less from reforms when they are old. Hence, they want to have more power about future reforms, which gives them an incentive to agree on a supermajority rule. Erlenmaier and Gersbach (2004) argue that first best outcomes can be achieved under unanimity, provided that it is supplemented by a number of constitutional provisions, such as bundling of projects. Compared to all these papers, the structure of our model is more parsimonious, in particular, as we are excluding side payments. In addition, we focus on the effects of inner group formation on the efficiency of an organization in the absence of constitutional rules, i.e., under voluntary cooperation.

Harstad (2006) investigates how flexible cooperation (organization members can decide on the speed of integration) compares to rigid cooperation (all members go at the same speed). While addressing similar issues, his model does not consider the disciplining role that a threat of an inner club has on weak members. In Dixit (2003) this role is played by network externalities. Owing to these externalities, agents may sequentially adopt an innovation (or join an organization) even though the introduction is not in their collective interest. That is, adoption is individually rational, unless agents can coordinate their actions. In our model, weaker members are in a similar situation – they would prefer the threat of forming an inner organization not to exist. In addition, stronger members can execute the threat and form an inner organization, a possibility not explored by Dixit.

The paper is organized as follows. Section 2 outlines and solves the basic model in which inner clubs are not an option. Section 3 introduces this possibility and examines the impact that threat of an inner club has on the initial organization. Section 4 derives the conditions under which an inner club forms and characterizes the equilibrium outcomes. Section 5 discusses key assumptions. Section 6 analyses organizations operating under majority rules. Section 7 discusses European Integration as an illustration of our theory. Concluding remarks are in Section 8. Formal proofs are relegated to the Appendix.

2 The Curse of Unanimity

We consider an organization with N members, who produce a common good.[3] The provision of the good increases in the size of the organization and in the effort e of the members. Inspired by Leontief partnership models (e.g., Vislie, 1994), we assume that the amount of the good is determined by the smallest effort in the organization, scaled by the size of the organization: $N \min[e_1, e_2, ...e_N]$.

The utility of each member increases in the consumption and decreases in effort. The benefit from consumption is the same for all members, whereas the effort cost differs across members. Member $i \in N$ has effort cost $\theta_i e^2/2$, and the type parameter θ_i is distributed on the support $[\underline{\theta}, \bar{\theta}]$. Furthermore, the productivity difference between any two adjacent members is the same. We refer to $\underline{\theta}$ as the most productive or "strongest" type, and to $\bar{\theta}$ as the least productive or "weakest" type. Assigning rank 1 to the strongest type $\underline{\theta}$, the cost parameter of the member with rank i is

$$\theta_i = \underline{\theta} + \frac{i-1}{N-1}\left(\bar{\theta} - \underline{\theta}\right). \tag{1}$$

Given that the good is produced with a Leontief technology, member i's payoff is

$$y(\theta_i, e) = N \min\{e_1, ..e_N\} - \theta_i e_i^2/2.$$

As the members have different costs, their preferred amount of common good differs. Hence, some members could offer side payments to others in order to influence their effort choices. However, we focus on the threat of forming a "club-in-the-club" as a mechanism to overcome the opposition of individual members against reform proposals. Therefore, we abstract from transfer payments.

Production of the good is modelled as a two-stage game. In the first stage, members vote on a minimum effort level in the club and in the second stage each member

[3] Our interest is how an existing organization responds to new challenges for which its members have different preferences. Hence, we abstract from the question of whether any given member has an incentive to leave the organization or whether outsiders would like to join.

simultaneously exerts an effort. Individual effort levels are verifiable and each member commits herself to exert - at least - the effort level agreed upon in the voting stage. That is, underprovision is infinitely punished, but the voting outcome is not binding from above. The asymmetry reflects our interest in the constraints that unanimity imposes on organizations. However, unilateral overprovision is never an equilibrium outcome due to Leontief technology.[4]

In standard voting procedures agents vote over pairs of alternatives and the winner in one round is posed against another alternative in the next round. Under the unanimity rule, this procedure may easily fail to generate a unique winner. Further, the outcome of the unanimity vote is highly sensitive to the order in which proposals are put to the vote, as well as the default option in case none of the alternatives receives unanimous support. That is, there is no robust unanimity voting procedure, and the literature has not agreed on a standard modelling approach.

Motivated by the interest in the impact that the weakest member has on the club production, we propose a procedure that parallels that of the continuous-time ascending-bid auction (Milgrom and Weber, 1982).[5] An uninterested agent ("auctioneer") proposes a sequence of continuously increasing effort levels $\{e\}$ starting with the initial level $e = 0$. After each proposal agents decide whether or not to vote in favour of a further increase in the common effort level. Once a member "leaves the auction" by voting against an increase, she cannot "return" by supporting any subsequent proposals. Under the unanimity rule, voting stops once a single member exits the vote. Accordingly, the option to withdraw from the voting gives veto power to each member. After the voting stage, members simultaneously choose their effort and the good is produced.

In this game, Nash equilibria are outcomes in which all members exert some common effort $e \in \left[0, N/\bar{\theta}\right]$, where $N/\bar{\theta}$ is the effort maximizing the payoff of the weakest type. More precisely, any effort $e \in \left[0, N/\bar{\theta}\right]$ can be supported in an equilibrium where at least two members withdraw from the vote at some $e^V \in [0, e]$ and where all members choose in the production stage the same effort level e. Indeed, given some e^V, member i's decision problem at the implementation stage is

$$\max_{e_i \geq e^V} \left(N \min\{e, e_i\} - \theta_i e_i^2/2 \right).$$

Member i's preferred choice $e_i^* = N/\theta_i$ exceeds e, as $e \leq N/\bar{\theta} \leq N/\theta_i$. Thus, member i always chooses $e_i = e$, since any effort $e_i - e > 0$ would be wasted. That is, unilateral

[4] Due to the Leontief technology, there is also no loss of generality in assuming that the organization votes on a common (minimum) effort level as opposed to a menu of type-contingent efforts.

[5] Our procedure is not robust either. For instance, a decending order would favor the productive members.

overperformance ($e_i > e$) is never profitable. Note that the voting outcome e^V needs not to be binding as all members can choose to exert higher effort $e \geq e^V$.

At the voting stage member i's only deviation that influences the outcome of the game is to withdraw prior to e^V. This deviation is profitable iff $e^V > N/\theta_i$. By withdrawing at $e_i^V \leq N/\theta_i$ and choosing $e_i^* = N/\theta_i$, member i attains her first best in the implementation stage. Since this applies to all members $i = 1, .., N$,

$$e^V \leq N/\bar{\theta} \tag{2}$$

must hold in equilibrium. Consequently, any effort $e > N/\bar{\theta}$ cannot be an equilibrium outcome. Indeed, if everyone but member N chooses e, member N's unilateral underperformance ($e_N^* = N/\bar{\theta} < e$) is both profitable and compatible with the voting outcome as $N/\bar{\theta} \geq e^V$ by (2).

It is well known that input games for a team with a Leontief technology have a continuum of Nash equilibria and that these equilibria can be Pareto-ranked.[6] This also holds for our voting game: all members prefer the Pareto-dominant equilibrium with $e = N/\bar{\theta}$ which we use as a benchmark in the subsequent analysis.

Proposition 1 *Under unanimity, the weakest member of the organization executes her veto power, holding back the entire organization at her privately optimal choice.*

Proposition 1 captures the idea that unanimity voting may result in the weakest member blocking any attempt to increase organization-wide effort. In principle, unanimity could well favour stronger rather than weaker members of an organization. For example, more productive (and wealthier) members would exercise their veto power if the organization were to vote on redistribution and not on effort. However, we follow the wide-spread view that unanimity tends to protect weak members and slow down reforms (e.g., Erlenmaier and Gersbach, 2004).

3 Undermining Veto Power

We now show how the veto power of weaker members can be undermined by the threat of some members to form an "inner organization". This threat may lead to three different types of outcomes: i) "initial organization", the equilibrium outcome with no inner organization and no additional effort; ii) "reformed" organization, the outcome with no inner organization but higher organization-wide; iii) "divided" organization,

[6] The inevitable free-riding problem in teams where the members' inputs are substitutes can be avoided when inputs are strict complements. For such teams, there exists a linear (balanced-budget) sharing rule that implements the efficient outcome as a Nash equilibrium outcome (Legros and Matthews, 1993; Vislie, 1994).

the outcome with a club-in-the-club. Here we analyze the first two outcome types and relegate the analysis of divided organizations to the next section.

Each member can freely decide whether she wants to join the inner club. That is, the constitution of the initial organization is incomplete. To keep the analysis tractable, we abstract from the possibility of multiple inner organizations and allow for at most one inner organization. Furthermore, the inner organization must have at least two members ($n \geq 2$). This is a natural restriction because an inner organization provides a public - rather than private - good to its members.

Instead of adopting a multi-task framework (Holmström and Milgrom, 1991) which would view inner and outer club efforts as substitutes, we assume a negative externality in consumption: An inner organization with n members reduces the utility of consuming the outer club good for all N agents by λn with $\lambda \geq 0$. The deadweight loss λ is meant to capture the notion that the formation of an inner club causes its members to divert attention and effort from the outer organization.[7]

For symmetry, the production technology of the inner organization is the same as the one of the outer organization. Membership in the inner organization generates additional per-capita benefits of $n(e_{In} - e_{Out})$, where e_{In} (e_{Out}) denotes the minimal effort exerted by anyone who is a member of the inner (outer) organization. Notice that members' efforts are still complements in the production functions of the inner and outer organization, but that an inner organization allows for the possibility to exert additional effort.

We use the term "constellation" for a partitioning of members into an inner organization with $n \leq N$ members, together with the associated effort levels in the outer and the inner organization. The payoff of type i who is a member of both the inner and the outer organization is

$$y_i = n(e_{In} - e_{Out}) + Ne_{Out} - \lambda n - \theta_i e_i^2/2.$$

The payoff of type j who is only a member of the outer organization is

$$y_j = Ne_{Out} - \lambda n - \theta_j e_j^2/2.$$

In general, the formation of an inner organization is sensitive to how agents coordinate, for instance who determines the effort level e_{In} and/or the size n of the inner club. We intentionally abstract from coordination mechanisms through some arbitrary agenda-setting procedure. Rather, we let Nature choose e_{In}. This allows us to identify all constellations that can be supported as Nash equilibrium outcomes. Endogenizing

[7] The type of externality associated with an inner club is not crucial for the analysis. For instance, we obtain qualitatively the same results in a setting with a (lump-sum) inner club membership benefit.

the choice of e_{In} poses both technical and severe conceptual problems. If the size of the inner club were given, one can think of one or the other procedure, that would generate e_{In} compatible with this size. However, there is no obvious rationale for selecting one inner club size over the other one. Also, there is no plausible justification for why the inner club size would be chosen prior to the amount of the inner club good, or vice versa. Last but not least, we strongly believe that any procedure that simultaneously determines n and e_{In} would be highly arbitrary. Therefore, we choose to characterize all Nash equilibrium outcomes. Besides being of interest *per se*, these constellations also constitute the constraints of any potential decision-maker's optimization problem.

The production of the outer and possible inner club goods takes place in three stages. In the first stage, members vote on the minimum effort of the outer organization. As before, voting follows the ascending procedure under the unanimity rule. In the second stage, Nature draws e_{In}. Following the logic of the model, we restrict the possible draws of nature to $e_{In} > e_{Out}$. Having observed e_{In}, all members have the option to simultaneously subscribe to join the inner club. By subscribing, a member commits to exert e_{In}. Otherwise, she gets infinitely punished. Non-subscribers cannot be members of the inner organization, irrespective of their subsequent effort choice. In the final stage, all members simultaneously choose their effort and the club goods are produced.

We make two further assumptions whose implications are discussed in Section 5. First, Nature's draw e_{In} is strictly binding in the sense that inner club members have to exert exactly e_{In}, neither less nor more. This simplification allows to avoid multiplicity of equilibria. Second, each member's decision to withdraw from the voting is non-strategic in the sense of ignoring its impact on the subsequent subscription decision of other members. This helps to keep the model tractable, though we argue in Section 5 that strategic voting would not change the qualitative results.

The assumption of sincere voting pins down a unique voting outcome, as each member i withdraws at her preferred effort level N/θ_i. So the weakest member ends the voting by exiting at $e^V = e_{Out} = N/\bar{\theta}$. That is, non-strategic voting rules out all Pareto-inferior equilibria of the basic framework.

If Nature draws a moderate level of e_{In}, there exists a Nash equilibrium where all members subscribe to join the inner organization and exert exactly e_{In}. Consider the choice of the weakest member when all other members subscribe to e_{In}. If she also subscribes, she exactly matches the announced threshold e_{In} as any other effort level entails an infinite penalty. Alternatively, if she abstains from joining, an inner club of size $N-1$ forms. She then sets the outer-club effort to her most preferred level $e_N^* = N/\bar{\theta}$. This option entails lower disutility of effort but also lower consumption

and in addition the deadweight loss $\lambda(N-1)$. Comparing the respective payoffs of the weakest type

$$Ne_{In} - \bar{\theta}e_{In}^2/2 \geq N(N/\bar{\theta}) - \lambda(N-1) - \bar{\theta}\left(N/\bar{\theta}\right)^2/2 \tag{3}$$

reveals that she prefers to join the inner organization for all efforts

$$e_{In} \in \left(N/\overline{\theta}, N/\overline{\theta} + \sqrt{2\lambda(N-1)/\overline{\theta}}\right].$$

Indeed, solving the quadratic inequality (3) for e_{In} and imposing the constraint $e_{In} > e_{Out}$ yields the above interval.

Suppose type N joins and consider type $N-1$. Similarly to the weakest type, if she does not join the inner organization, she would want to set the outer-club effort to her preferred level $e_{N-1}^* = N/\theta_{N-1}$. If $e_{In} > N/\theta_{N-1}$ this deviation cannot be profitable because type $N-1$ has to exert less "additional" effort $(e_{In} - e_{N-1}^*)$ at a lower cost than the weakest type who still prefers to join. If $e_{In} < N/\theta_{N-1}$ her preferred effort level is not even feasible as the outer-club effort cannot exceed the inner-club one. This follows from the Leontief technology and the fact that members of the inner club continue to be members of the outer club. Thus, type $N-1$ also prefers to subscribe to the inner organization. This reasoning applies to all other types $i = 1, .., N-2$. Consequently, there exists an equilibrium with all N members choosing the same effort level e_{In} as long as e_{In} does not exceed $e^{RO} = N/\overline{\theta} + \sqrt{2\lambda(N-1)/\overline{\theta}}$.

Proposition 2 *Reformed organizations can emerge for any $\lambda > 0$ and $e_{In} \in \left(N/\overline{\theta}, e^{RO}\right]$. That is, the threat of forming an inner club undermines the veto power of the weakest member and increases organization-wide effort.*

Unanimity is commonly viewed as preventing majorities from coercing minorities at the cost of organizational inertia or inability of adjusting. Proposition 2 shows that this view needs to be qualified: Unanimity need not be tantamount to complete protection of the weakest members or, equivalently, to the inability to reform. The threat of forming an inner organization can undermine the veto power of each single member and may enable the organization to reform. To be an effective reform mechanism, two conditions must hold. First, the statutes of the organization must be incomplete, thereby exempting the formation of an inner organization from the unanimous approval. Otherwise, weaker members would have no reason to avoid the formation of an inner organization by exerting more effort. Rather, they could simply veto its formation. Second, the inner organization must impose some externalities on the outer organization. Otherwise, the weaker members have no incentives to increase

their effort beyond their privately optimal level. The "reform potential" of an organization, measured by the difference $e^{RO} - N/\bar{\theta}$, increases with the deadweight loss associated with an inner organization and with the size of the initial organization.

4 Club-in-the-club

We have so far only looked at equilibrium constellations in which there are no inner organizations in equilibrium. We here explore the set of constellations with divided organizations that can be supported as Nash equilibrium outcomes for any given deadweight loss λ.

Assume an inner organization exists. Due to non-strategic voting, all agents who do not subscribe to the inner club exert the effort level $e_{Out} = e^V = N/\bar{\theta}$. As the inner club effort level e_{In}, drawn by Nature, is binding, all members who join the inner organization exert e_{In}.

For expositional simplicity, we only consider inner clubs with $n \in \{3, ..., N-2\}$, where the relevant participation constraints have the same functional form. While our setting allows for $n = 2$ or $n = N - 1$, the respective participation constraints differ slightly in these two cases (see appendix for details). Occasionally, we will comment on the latter cases, but if not explicitly stated, our discussion refers to $n \in \{3, ..., N-2\}$.

We first establish which types are members of both the inner and outer organization.

Lemma 1 *Provided that an inner organization of size $n \in \{3, ..., N-2\}$ is an equilibrium outcome, its members are the low ranked (most productive) types $i \in \{1, ..., n\}$.*

For an inner organization of size n to exist, two types of constraints must be satisfied. First, $N - n$ members of the outer organization must prefer staying in the outer organization rather than joining the inner organization. Second, the n members must prefer to be in the inner organization.

An agent i chooses to be a member of the inner organization if the following condition holds:

$$n(e_{In} - e_{Out}) + Ne_{Out} - \lambda n - \theta_i e_{In}^2/2 \geq Ne_{Out} - \lambda(n-1) - \theta_i e_{Out}^2/2.$$

Rearranging yields

$$n(e_{In} - e_{out}) - \lambda \geq \theta_i \left[e_{In}^2 - e_{Out}^2\right]/2\,. \tag{4}$$

The LHS of this constraint is composed of the per-capita membership benefits of the inner organization net of the deadweight loss λ. Both are independent of the type

i. The RHS is type i's cost differential between exerting the inner and outer effort levels and increases in i. Thus, if condition (4) holds for type i, it must hold for all more productive types $j = 1, ..., i-1$. Hence, given the equilibrium inner club has n members, these members must be exactly the n most productive agents.

The above result implies that an inner organization forms in equilibrium if the following two constraints are satisfied:

$$n\left(e_{In} - N/\bar{\theta}\right) - \lambda \geq \theta_n \left[e_{In}^2 - (N/\bar{\theta})^2\right]/2\,, \tag{5}$$

$$(n+1)\left(e_{In} - N/\bar{\theta}\right) - \lambda < \theta_{n+1} \left[e_{In}^2 - (N/\bar{\theta})^2\right]/2\,. \tag{6}$$

The first condition ensures that the marginal, i.e. least productive, member of the inner organization prefers to be member in both the inner and the outer organization. The second condition ensures that the most productive member in the outer organization prefers to be member in the outer organization only. We assume that type $n+1$ does not join in case she is indifferent, which accounts for the strict inequality in the non-participation constraint (6).

We now establish the conditions for the existence of divided organizations and then characterize constellations supporting inner organizations of different size.

An increase in the size of the inner organization benefits all its members as it rises the amount of the inner public good. However, the less productive agents may find it too costly to exert the requested effort level e_{In}. Hence, any inner organization strikes a balance between size and productivity of its marginal member. This trade-off has no interior solution if the productivity differences among (two adjacent) members is relatively small. That is, when the N members are relatively homogeneous, an inner organization never forms in equilibrium.

Proposition 3 *An inner club can only emerge if agents are sufficiently heterogeneous,* $N < \bar{\theta}/\underline{\theta}$.

When members are relatively homogenous, the benefit of increasing the size of the inner organization exceeds the (possibly) adverse effect on the effort levels that are compatible with the resulting inner organization. As shown in Lemma A1 in the Appendix, for $N > \bar{\theta}/\underline{\theta}$ the most preferred effort level of the marginal member of the inner club, $e_n^* = n/\theta_n$, increases in n, the size of the inner club.[8] As e_{In} drawn by Nature exceeds $N/\bar{\theta}$, no marginal member would ever subscribe to e_{In}. Indeed, her alternative is to exert the $e_{Out} = N/\bar{\theta}$, which by Lemma A1 differs less from her most preferred inner club effort level, and does not entail an externality.

[8] Given that $y_i = n(e_{In} - N/\bar{\theta}) + N(N/\bar{\theta}) - \lambda n - \theta_i e_i^2/2$ is the payoff of type i who is a member of the inner organization of size n and the outer organization, type i would choose $e_{In} = e_i^*(n) = n/\theta_i$.

For the remainder of this section, we assume that the heterogeneity condition $N < \overline{\theta}/\underline{\theta}$ holds. This is, however, only a necessary condition for the existence of inner organizations. Rather intuitively, the size of the deadweight loss and the level of the inner club effort, drawn by nature, also matter. Indeed, even when members are heterogeneous, divided organizations only exist for certain pairs of (e_{In}, λ).

Denote by Ω the set of all pairs (e_{In}, λ) that satisfy the two inequalities

$$3\left(e_{In} - N/\bar{\theta}\right) - \lambda \geq \theta_3\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2\,, \tag{7}$$

$$(N-1)\left(e_{In} - N/\bar{\theta}\right) - \lambda < \theta_{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2\,. \tag{8}$$

where (7) is type 3's participation constraint and (8) is type $N-1$'s non-participation constraint. These constraints define the largest set of pairs (e_{In}, λ) that support a divided organization outcome.

Indeed, sufficient heterogeneity implies that the effort cost θ_n increases faster than n. Hence, if the non-participation constraint

$$n(e_{In} - e_{out}) - \lambda < \theta_n\left[e_{In}^2 - e_{Out}^2\right]/2.$$

holds for $n = 3$, it holds for all less productive types $k > 3$. That is, if type 3 does not want to be the marginal member of the inner club of size 3, no type $k > 3$ wants to be the marginal member of the inner club of size k. At the same time, if type n wants to be the marginal member of the inner club of size n, any more productive type $k < n$ has the same preference. Finally, member $N-1$'s non-participation constraint has to be met, as we only consider inner clubs of size $n \in \{3, .., N-2\}$.

Proposition 4 *Provided that types are sufficiently heterogenous, an inner club of size $n \in \{3, .., N-2\}$ can form in equilibrium iff the pair (e_{In}, λ) belongs to the set Ω. Moreover, for each pair (e_{In}, λ) the size and composition of the inner club is unique.*

As was just pointed out, type 3's willingness to join is crucial for the existence of a divided organization of any size. Type 3 refrains from subscribing to the inner club, if the deadweight loss λ is very high. Signing up for very high effort level e_{In} is too costly for type 3, so, again, she stays in the initial organization. Also, if the inner club effort level is not very different from the outer club one, there is little value for type 3 in joining the inner club and suffering the dead-weight loss λ.

Figure 1 depicts the set of pairs (λ, e_{In}) that can support divided organizations. Drawn in the λ-e_{In} space, the boundaries of this set are two parabolas intercepting the vertical axis. The outer parabola is determined by the participation constraint of type 3 and the inner parabola by the non-participation constraint of type $N-1$.

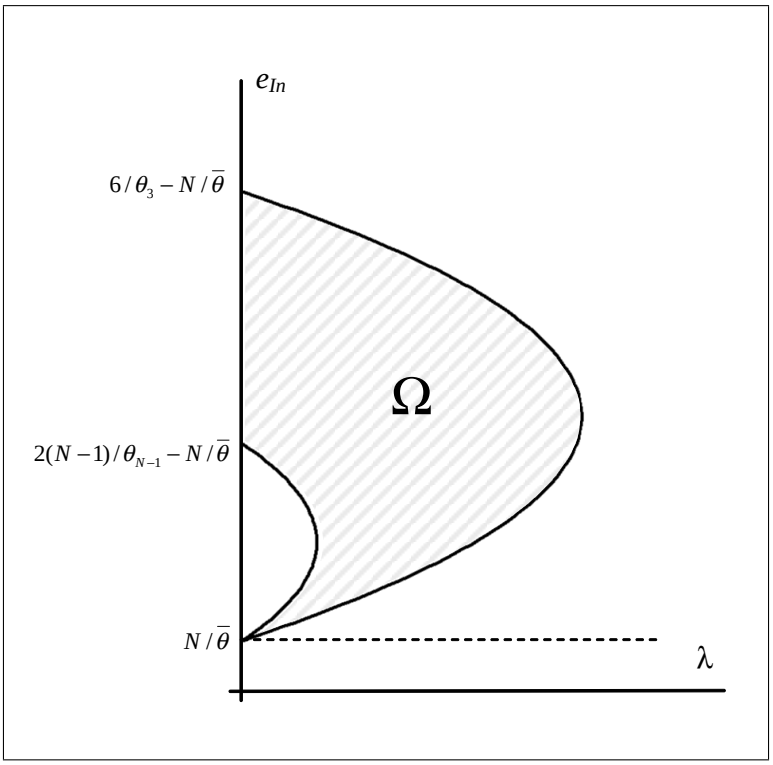

Figure 1.

Setting $\lambda = 0$ in the participation constraints and rearranging yields the respective e_{In} intercepts.

The uniqueness is easily understood by comparing the conditions for the existence of two inner organizations that differ in size by one member. On the one hand, type k prefers to be a member of the inner organization of size k if

$$k\left(e_{In} - N/\bar{\theta}\right) - \lambda \geq \theta_k \left[e_{In}^2 - (N/\bar{\theta})^2\right]/2$$

(condition 5) holds. On the other hand, type k prefers not to join the inner organization of size $k-1$ if

$$k\left(e_{In} - N/\bar{\theta}\right) - \lambda < \theta_k \left[e_{In}^2 - (N/\bar{\theta})^2\right]/2$$

(condition 6) holds. The inequality sign apart, these two constraints are identical. Accordingly, there exists no effort level e_{In} that can be supported for a given deadweight loss by more than one inner organization.

More generally, heterogeneity implies that if type k does not want to be the marginal member of an inner organization of size k, neither of types $k+1, k+2, ...$ wants to be the marginal member of the inner club of respective size. Therefore, inner clubs with non-adjacent size cannot coexist. In addition, the unique composition of the club

follows from Lemma 1, which establishes that an inner organization of size n consists of the n most productive members.

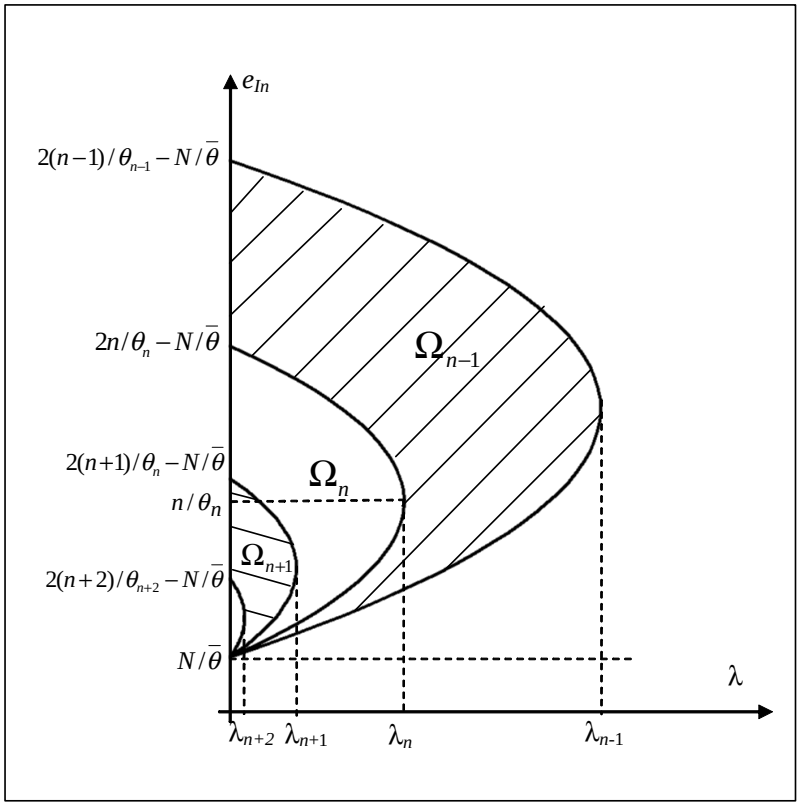

Figure 2.

The uniqueness result implies that the set Ω can be partitioned into subsets Ω_n, each corresponding to all pairs (λ, e_{In}) consistent with an inner organization of size n. Figure 2 plots the subsets Ω_n in the λ-e_{In} space. These subsets have an "onion-like" shape with the outer layers enclosing the pairs compatible with smaller inner organizations. Each subset Ω_n is determined by type n's participation constraint and type $n+1$'s non-participation constraint. The formal characterization is provided in Appendix A.6. When the deadweight loss is relatively large ($\lambda > \lambda_n$), type n strictly prefers not joining the inner organization, even if nature drew type n's preferred effort level, $e_n^*(n) = n/\theta_n$. For intermediate values of the deadweight loss ($\lambda \in (\lambda_{n+1}, \lambda_n]$), there exist values of e_{In} such that type n's benefits of being in the inner organization exceed the deadweight loss λ. For small λ the range of e_{In} for which type n prefers to be part of the inner organization becomes larger and includes values for which type $n+1$ would also want to join the inner organization. These values are centered around type $n+1$'s best preferred effort level, $e_{n+1}^*(n+1) = (n+1)/\theta_{n+1}$. To restore the non-participation constraint of type $n+1$ (condition 6), these values of e_{In} have to be

removed from the Ω_n set. Consequently, the solutions to conditions (5) and (6) are two disjoint intervals when $\lambda_{n+1} \geq \lambda > 0$.

As the discussion above indicates, the maximum size of an inner organization that can be sustained decreases (weakly) in the deadweight loss λ. Intuitively, for high λ the difference in the effort costs between the marginal inner club member θ_n and the least productive type $\overline{\theta}$ needs to be sufficiently high, otherwise n stays outside. For a given type distribution it means that the size of the maximum supportable inner club shrinks as λ increases.

To complete the analysis we address the coexistence of initial, reformed and divided organizations. For any parameter values, the initial organization is an equilibrium of the game. This follows from the assumption that an inner organization must have at least two members. Hence, if all other members choose not to subscribe to e_{In} and exert the weakest member's preferred effort level, no single member has an incentive to deviate from this common pattern. In addition, if the effort e_{In} is not too high, the reformed organization can exist in equilibrium. The divided organization requires sufficiently heterogeneous members in combination with moderate levels of e_{In} and deadweight loss λ.

Proposition 5 *If the pair (e_{In}, λ) belongs to Ω, the heterogeneity condition holds and $e_{In} \leq e^{RO}$, all three organization outcomes coexist.*

More specifically, a pair $\left\{(e_{In}, \lambda) \in \Omega : e_{In} \leq e^{RO}\right\}$ can result both in an equilibrium with a unique inner club and a reformed organization with all N members exerting e_{In}. Clearly, the initial organization is also supported. The coexistence is due to the fact that an inner club gives rise to a deadweight loss. The least efficient member may prefer to stay outside if at least one more member stays outside (that is, $n \leq N-2$). In this case joining the inner club has double cost: it involves exerting a higher effort e_{In} and bearing a higher deadweight loss. Instead, if the other $N-1$ types exert e_{In}, by joining them the least efficient type faces the cost of exerting too high effort but at the same time eliminates the inner club externality. Thus, the weakest member may choose to join and exert e_{In} if all the other types join. Figure 3 depicts the coexistence region of reformed and divided organization.[9]

[9] This does not hold for the inner organization of size $N-1$, as member N's decision to exert either e_{In} or $N/\bar{\theta}$ fully determines the equilibrium constellation. That is, divided and reformed organization cannot co-exist. This constitutes the only qualitative difference of the inner club of size $N-1$ relative to all smaller inner clubs.

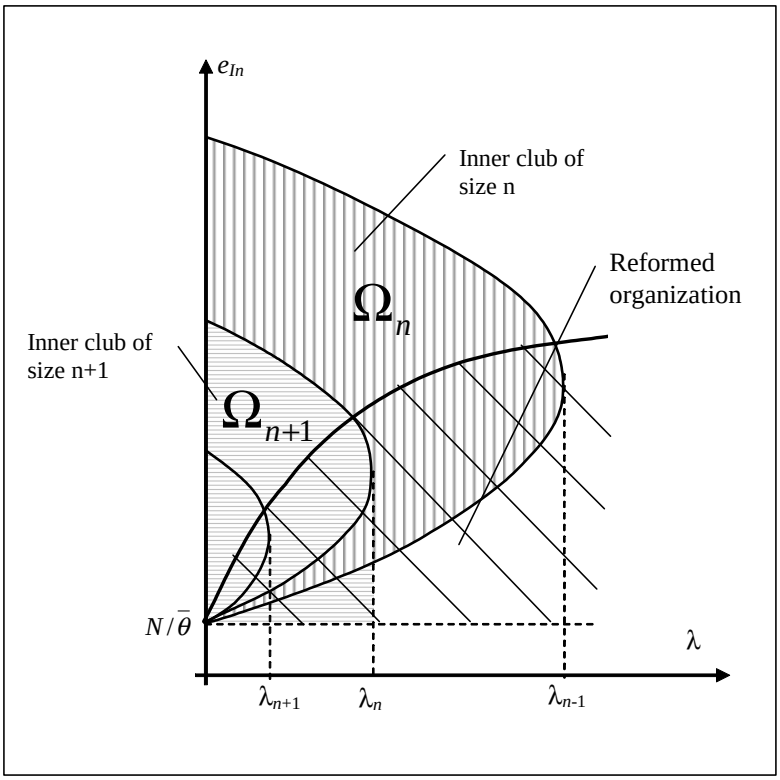

Figure 3.

5 Discussion

Throughout the analysis, we rely on several core assumptions to keep the model tractable. We now discuss their implications for the results. The assumption of equidistantly distributed types delivers a generic functional form for the participation constraints of the inner club members. The essential feature ensuring the formation of divided organization is the heterogeneity of types, that is, the most preferred effort of the marginal inner club member $e_n^*(n) = n/\theta_n$ is decreasing in the club size. We are confident that any distribution satisfying this property can generate divided organization equilibria. The "thickness" of the Ω_n-layers would, however, differ as compared to the equidistant distribution. The distribution has no impact on the formation of the initial and reformed organizations in equilibrium, as they are solely determined by the decisions of the least productive member.

In the model, Nature's draw e_{In} is assumed to be binding not only from below but - unlike in the voting stage - also from above. This simplifying assumption ensures that the members of the inner club exert precisely e_{In} which for a given deadweight loss λ yields a unique divided organization. If inner club members were free to exert a higher effort than what Nature draws, inner clubs of different size could emerge in

equilibrium. Suppose that a pair (e_{In}, λ) supports an inner club of size n. Then, by coordinating to work harder than e_{In}, the most productive $m < n$ members can form an inner organization in equilibrium. To see this, consider a point $(\tilde{e}_{In}, \tilde{\lambda}) \in \Omega_n$ in Figure 2. The ray along the vertical line $\lambda = \tilde{\lambda}$ starting at $\tilde{e}_{In}$ and corresponding to an increase in e_{In}, crosses all the sets Ω_{n-1},Ω_{n-2},...,Ω_3. Setting the inner club effort to equal exactly the Nature drawn level allows us to convey our ideas, while keeping the analysis tractable.

While we restrict our analysis to a single inner club, the logic of our model seems compatible with multiple inner organizations. For example, if Nature draws two e_{In}, a plausible equilibrium candidate is a constellation with two inner clubs, the most productive types being members of both inner clubs, intermediate types joining the "outer-inner" club and the least productive types being only in the outer organization. However, the outcomes in such an extended framework will depend on modelling details such as the assumed interaction between the deadweight loss of different inner clubs, and single vs. multiple inner club membership.

Finally, we turn to the assumption of non-strategic voting. While it is a standard assumption in many political economy models, it may be limiting in our framework. Indeed, the prospect of an inner club provides the members of the initial organization with the incentive to behave strategically in the voting stage. On the one hand, more productive types may choose to withdraw from the voting before the least productive member would pull out when voting sincerely. While this reduces the provision of the outer club good, it may induce more members to subscribe to the emerging inner club. On the other hand, less productive members may remain in the voting beyond their most preferred level. Though costly, the extra effort reduces the attractiveness of an inner club, thereby lowering the number of its potential members and the consequent deadweight loss, or even preventing its formation altogether.

Therefore, allowing for strategic voting would likely alter the effort level of the outer club and the size of the inner organization. Nonetheless, we would expect to observe the same types of organizational outcomes: divided as well as reformed and initial organizations. In addition, the game may feature an equilibrium in which all members exert an effort below the most preferred level of the least productive member.

It is worth noting that strategic voting entails certain costs but uncertain benefits. When a highly productive member withdraws early, the outer club good is provided at the lower level. At the same time, a larger inner club may or may not materialize depending on the draw of e_{In}. More generally, the benefits of strategic voting depend on the extent to which the agent can influence or correctly anticipate the subsequent decision (i.e., the level e_{In}). Our setting abstracts from any specific agenda setting

procedure and lets e_{In} be randomly chosen by Nature. In this complex environment, the benefits of strategic voting seem particularly limited, making the sincere voting assumption less restrictive than it may seem at first glance.

6 Majority Rules

In divided organization outcome, weaker members are not forced to provide more effort than their privately optimal choice. Thus, unanimity protects weak members from the tyranny of the majority but at the price of the formation of a club-in-the-club. Many clubs may want to avoid becoming a two-class organization. One possible remedy is a majority rule since it limits the decision power of the weak members. This reduces the extent to which more productive members are held back which in turn may prevent the formation of inner clubs.

We now consider organizations operating under different majority rules $\mathcal{M}(m)$, where the majority threshold $m \in [0.5, 1)$ corresponds to the required fraction of supporting votes. As before, voting follows the ascending procedure, but under the $\mathcal{M}(m)$ majority rule it ends once a fraction $(1-m)$ of agents has chosen to "leave the auction".[10] In the second stage Nature draws a (potential) inner club effort that exceeds the one voted upon in the first stage.

Under majority rule $\mathcal{M}(m)$ the organization-wide effort e^m_{Out} coincides with the best-preferred choice of its decisive member mN. For instance, the median type $0.5N$ is decisive in case of the simple majority rule $\mathcal{M}(0.5)$, and the resulting effort is $e^{0.5}_{Out} = N/\theta_{0.5N}$. As the majority threshold m increases the decisive member mN becomes less productive and the organization-wide effort e^m_{Out} declines. Hence, stronger members have more incentives to form an inner organization. If majority rules are at all effective in preventing inner clubs, they must have sufficiently low thresholds.[11]

Proposition 6 *Under the simple majority rule a divided organization never emerges.*

The high organization-wide effort level under the simple majority rule $\mathcal{M}(0.5)$ makes inner clubs no longer attractive even for the most productive members. That is, there are no pairs (e_{in}, λ) that support the formation of an inner organization of any size.

While the simple majority rule succeeds in preventing inner clubs, it leaves weaker members without protection against the tyranny of the majority. Clearly, a superma-

[10] For notational simplicity, we abstract from the integer problem.

[11] For the same reason as under unanimity, initial organization with all members exerting e^m_{Out} is always an an equilibrium outcome under the majority rule $\mathcal{M}(m)$. Similarly, one can also support reformed organization equilibria.

jority rule would coerce weaker members less. But can it still preclude the formation of an inner organization? The analysis so far has shown that divided organizations form under unanimity rule ($m = 1$), provided agents are heterogenous. Proposition (6) establishes the inexistence of inner club equilibria under the simple majority rule ($m = 0.5$). Based on a continuity argument one may expect this result to be obtained already under a *qualified* majority rule.

Proposition 7 *For any initial organization* $(\underline{\theta}, \overline{\theta}, N)$ *there exists a majority threshold* $\overline{m}(\underline{\theta}, \overline{\theta}, N) > 1/2$ *such that no inner organization emerges under all majority rules* $\mathcal{M}(m)$ *with* $m < \overline{m}(\underline{\theta}, \overline{\theta}, N)$. *The threshold* $\overline{m}(\underline{\theta}, \overline{\theta}, N)$ *decreases as agents become more heterogeneous.*

The exact majority threshold depends on the characteristics of the organization. When an organization is more heterogeneous, as measured by an increase in $\overline{\theta}$ (decrease in $\underline{\theta}$), the incentives of its members are less aligned. Under a given majority rule, productive members are held back to a larger extent, which makes them more eager to form an inner organization. This tendency can be counteracted by a lower majority threshold. It increases the productivity of the decisive club member and the organization-wide effort level, thereby eliminating incentives to form an inner club.

7 European Integration

The evolution of the European Union (EU) provides fitting examples of our theory. To map the model into EU reality, important concepts are: the benefit of the public good, the "effort" of the members, and the heterogeneity of the costs associated with this effort.

The benefit of the public good "European integration" has many faces. Some examples are the formation of the 1951 European Coal and Steel Community in which France, the Benelux countries, Italy and Germany coordinated their actions in these industries of high military importance. In the aftermath of WW II, European integration was to bring about cooperation and to assure peace. On March 25, 1957, the six countries signed the Treaty of Rome creating the European Economic Community (EEC) with a view to promote trade among its member states. Further public goods were the creation of a single currency with its reduction of transaction costs in intra-European trade and the further extension of the integrated market through various enlargement waves.

The concept of effort in our theory has also multiple interpretations. For instance, to reap the benefits of European integration, countries must go through a number of

adjustment processes that take the time of politicians and bureaucrats, but also impose costs on the population. Laws must be changed and harmonized; languages must be learnt; opening markets exposes firms and workers to more competition. Reaching the Maastricht criteria in particular, committed national and subnational bodies alike to budgetary austerity, often with massive consequences for the population. Probably most important is the loss of sovereignty. This is a severe concern as the referenda and current discussion about the Lisbon treaty show: countries like Ireland, Poland, the Czech Republic or Germany have not yet ratified the treaty, because the public and parliament alike are concerned about sovereignty in general (the recent debate in Germany is about the role of national parliaments in EU integration decisions) or quite specific questions, such as abortion law in Ireland.

Heterogeneity between members can be treated in the model in two different ways. One could consider heterogeneity in the value associated with European integration, or as we do in the model, one can map heterogeneity into the cost function. The modeling strategies give similar results; in reality it is not straightforward to distinguish whether one or the other would be the source of heterogeneity. An early example is the plan for integration into a European Defence Community (EDC) in 1954. The French Parliament objected ratification and thus vetoed further integration. Whether France valued common defense lower than other members or estimated the costs (the potential loss of sovereignty) higher than others, seems a question that is secondary to our model. What is important, though, is to see the heterogeneity across countries in terms of the net benefits of integration.

Beyond justifying the structure of the model, it is also important to see to what extent outcomes of our model are in line with the reality of European integration. The failure of the European Defense Community is an early example of reform efforts that got vetoed by a member. The EU then saw many blocked reforms, but during the second half of 1980s, European Commission President Jacques Delors and some of the governments of stronger member states pushed for further integration. This process resulted in the Treaty of Maastricht, which states in article 2: "This Treaty marks a new stage in the process of creating an ever closer union among the peoples of Europe."

The core proposal to re-vitalize the EU was the creation of a common currency area with strict criteria for joining the "club-in-the-club", the European Monetary Union. Reaching the Maastricht criteria on public debt, deficit, interest rates and inflation meant to undertake efforts for each of the aspiring membership candidates. Naturally, these efforts would be more painful for countries with larger budgetary problems, such as Belgium, Greece or Italy. However, the benefits of further integration and the

creation of a joint currency would accrue to all participating members.

Arguably, the process of reaching the criteria led to a revitalization of the European integration process and a phase of growth. In the language of our model a group of economically stronger countries brought forward a proposal that was open to everyone. However, inclusion in the new club Euroland was only possible after exerting substantial efforts. The threat of forming such an inner club that would have excluded the underperformers seems to have worked. The countries that wanted to join managed to reach the criteria.

Our model also predicts that the risk of club-in-the club formation increase when heterogeneity of members increases and that a move from unanimity to qualified majority can be a remedy.

Indeed, the initial members of the European Community had quite similar aims and economic structures. Through a number of subsequent enlargement waves, the economic heterogeneity of EU members increased, thereby altering matters considerably. With the southern periphery joining, the challenge of keeping the new Union together had to be confronted as the size and use of structural funds, the state of labor markets and public administration provided ample reasons for conflict.

The Single European Act of Luxembourg (1986) can be seen as a first mild response. Here, unanimity was abandoned for many policy issues. This voting reform substantially reduced each single member's veto power. Despite such reforms, growing concerns about paralysis in the EU have spurred discussions about a two-speed Europe. Representatives of the stronger founding members, France and Germany (President Chirac and Former Foreign Minister Joschka Fischer) proposed to allow a subset of EU members to cooperate and integrate more. As in our model, larger heterogeneity increases the likelihood that inner clubs may form. As a response, the summit in Nice in 2000 explicitly set out to address the institutional problems associated with enlargement by re-weighting the allocation of votes in the Council and by extending qualified majority voting to an even larger number of areas. The 2001 intergovernmental conference in Nice was supposed to facilitate decision-making in the new larger Union and by regulating the formation of inner clubs through the instrument of "enhanced cooperation" among members. The Reform Treaty of 2007 regulates further the instrument of enhanced cooperation among sub-groups of countries, and reinforced the sole right of the Commission to formally propose such initiatives. The Reform Treaty also redefines qualified majority voting into double majority voting whereby a minimum of 55 per cent of Member States representing a minimum of 65 per cent of EU's population are required to pass legislation.

Thus, the dynamic of the European Union's voting system is well in line with the

logic of our theory in which the majority thresholds decline in the heterogeneity of club members.

8 Concluding Remarks

The paper presents a theory of loosely-knit organizations. While members have a common interest, there is no governance mechanism in place that enforces contributions to the common good. Hence, organization-wide decisions must be taken unanimously, granting each member veto power. We show that there are nonetheless ways for such organizations to avoid being held back by their least committed members. The threat of forming a club-in-the-club can induce members that are less interested or less productive to contribute more to the common good than privately preferred. Key for this mechanism is that the formation of a club-in-the-club imposes a deadweight loss on all members, but benefits only those who join the inner club. Then, unanimity does not preclude reform, in the sense of all members exerting more effort than is preferred by its weakest members.

We also show that identical organizations can end up quite differently: some may stagnate at the level preferred by its weakest members, others may reform, and yet others may be divided by the formation of an inner club. Furthermore, the divided organization outcome is more likely, when members are more heterogenous. To avoid this outcome an organization can adopt a majority rule. This constitutional change results in a higher organization-wide effort, and thus, often precludes the formation of an inner club. The change can be interpreted as a way of institutionalizing the reformed organization outcome, feasible under unanimity.

We illustrate our theory by the process of European integration: the introduction of the Euro has worked very much like a threat of an inner organization. While the EMU did not leave anyone behind, except those countries that decided to opt out, the increasing heterogeneity of EU members states creates the risk of a two-speed Europe. In response, the EU proceeds with putting more structure on enhanced cooperations and moving to qualified majority voting.

We believe that there are many more applications of our theory: the dynamics of research centres often exhibits similar tensions between individuals who are more or less committed to research, which often leads to infighting and the creation of sub-research centres. Many sport leagues suffer from a similar tension between high performance teams and those that lag behind, and there has been a threat of top teams to create their own superleagues, be it in basketball or football.

Our paper is only a first step to a more systematic analysis of loosely knit orga-

nizations and the club-in-the-club phenomenon. Strategic voting and the possibility of multiple competing inner clubs are extensions that we believe to be particularly interesting.

References

[1] Aghion, P., Bolton, P., 2003. Incomplete Social Contracts. Journal of the European Economic Association, 1(1): 38-67.

[2] Bolton, P., Roland, G., 1997. The Breakup of Nations: A Political Economy Analysis. Quarterly Journal of Economics, 112(4): 1057–90.

[3] Bordignon, M., Brusco, S., 2001. Optimal Secession Rules. European Economic Review, 45(10): 1811-1834.

[4] Buchanan, J., Faith, R. L., 1987. Secession and the Limits of Taxation: Toward a Theory of Internal Exit. American Economic Review, 77(5): 1023-1031.

[5] Cornes, R., Sandler, T., 1996. The Theory of Externalities, Public Goods, and Club Goods. Cambridge: Cambridge University Press.

[6] Dixit, A., 2003. Clubs with Entrapment. American Economic Review, 93(5): 1824-1829.

[7] Erlenmaier, U., Gersbach, H., 2004. Bundling and the Unanimity Rule. CEPR Discussion Paper 4808.

[8] Gradstein, M., 2004. Political Bargaining in a Federation: Buchanan meets Coase. European Economic Review, 48(5): 983-999.

[9] Harstad, B., 2006. Flexible Integration? Mandatory and Minimum Participation Rules. Scandinavian Journal of Economics, 108(4): 683-702.

[10] Holmström, B., Paul Milgrom, P., 1991. Multi-Task Principal-Agent Problems: Incentive Contracts, Asset Ownership, and Job Design. Journal of Law, Economics and Organization, 7(1): 24-52

[11] Legros, P., Matthews, S. A., 1993. Efficient and Nearly Efficient Partnerships. Review of Economic Studies, 60(3): 599-611.

[12] Messner, M., Polborn, M., 2004. Voting on Majority Rules. Review of Economic Studies, 71(1): 115-132.

[13] Milgrom, P.R., Roberts, J., 1992. Economics, Organization and Management. Prentice-Hall International, Englewood Cliffs, New Jersey.

[14] Milgrom, P. R., Weber, R. J., 1982. A Theory of Auctions and Competitive Bidding. Econometrica, 50(5): 1089-1122.

[15] Rousseau, J.-J., 1762. Contract Social, translated by Maurice Cranston. Penguin Books, London 1986.

[16] Vislie, J., 1994. Efficiency and Equilibria in Complementary Teams, Journal of Economic Behavior and Organization. 23(1): 83-91.

A APPENDIX

A.1 Inner clubs of size n=2 and n=N-1

For the inner clubs of size $n \in 3, ..., N-2$ the participation constraints (5) and (6) only differ with respect to the size of the inner club and the marginal members' productivity. For the inner club with $n = N - 1$ members, the non-participation constraint of type $n + 1$ is different. If type N were to join the inner organization, all members would exert the same effort and an inner organization would cease to exist. Thus, the non-participation constraint of type N is

$$\bar{\theta}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2 - (N-1)\lambda > N(e_{In} - N/\bar{\theta}).$$

A similar effect appears in the case of the inner club of size $n = 2$. As we do not allow for inner clubs consisting of one member, if type 2 does not join, the inner club fails to form. This is reflected in type 2's participation constraint

$$2\left(e_{In} - N/\bar{\theta}\right) - 2\lambda \geq \theta_2\left[e_{In}^2 - (N/\bar{\theta})^2\right]2.$$

These modified constraints do not substantially change the analysis, but they lead to different functional forms of the set of equilibrium effort level e_{In}.

A.2 Heterogeneity and Optimal Effort Choice

Given an inner organization of size n exists,

$$e_n^*(n) = \arg\max\left\{n(e_n - N/\bar{\theta}) + N^2/\bar{\theta} - \lambda n - \theta_n e_n^2/2\right\} = n/\theta_n$$

Lemma A1: *For $N < \bar{\theta}/\underline{\theta}$, n/θ_n increases with n. Otherwise n/θ_n decreases with n.*

Proof. Subtracting $(n+1)/\theta_{n+1}$ from n/θ_n yields

$$\frac{n\theta_{n+1} - (n+1)\theta_n}{\theta_n\theta_{n+1}} = \frac{\left[(n+1)(\theta_{n+1} - \theta_n) - \theta_{n+1}\right]}{\theta_n\theta_{n+1}}.$$

Using the definition $\theta_{n+1} = \underline{\theta} + n(\bar{\theta} - \underline{\theta})/(N-1)$ and $\theta_n - \theta_{n+1} = (\bar{\theta} - \underline{\theta})/(N-1)$ we obtain

$$\frac{1}{(N-1)\theta_n\theta_{n+1}}\left[(n+1)(\bar{\theta} - \underline{\theta}) - (N-1)\underline{\theta} - n(\bar{\theta} - \underline{\theta})\right]$$

$$= \frac{1}{(N-1)\theta_n\theta_{n+1}}\left[\bar{\theta} - N\underline{\theta}\right] > 0, \text{ if and only if } N\underline{\theta} - \bar{\theta} < 0.$$

■

A.3 Proof of Proposition 3

Proof by contradiction. If an equilibrium with an inner organization of size $n < N$ exists,

$$n(e_{In} - N/\bar{\theta}) - \lambda > \theta_n \left[e_{In}^2 - \left[N/\bar{\theta}\right]^2\right]/2$$

and

$$e_{In} > e_{Out}$$

must hold. Setting $e_{In} = N/\bar{\theta} + \delta$ and inserting it in the first condition yields

$$\delta(n - N\frac{\theta_n}{\bar{\theta}}) > \frac{\theta_n \delta^2}{2} + \lambda.$$

This can only hold if $(n - N\theta_n/\bar{\theta}) > 0$ or, equivalently, $(n/\theta_n - N/\bar{\theta}) > 0$. Using the definition

$$\theta_n = \frac{1}{N-1}\left[(N-n)\underline{\theta} + (n-1)\overline{\theta}\right],$$

the difference $(n/\theta_n - N/\bar{\theta})$ can be written as

$$\frac{(N-1)n}{\left[(N-n)\underline{\theta} + (n-1)\overline{\theta}\right]} - \frac{N}{\bar{\theta}} > 0.$$

Rearranging yields

$$\frac{(N-n)}{\bar{\theta}\left[(N-n)\underline{\theta} + (n-1)\overline{\theta}\right]}\left(\bar{\theta} - N\underline{\theta}\right) > 0$$

which contradicts $N \geq \overline{\theta}/\underline{\theta}$.

A.4 Characterization of the Ω Set

The Ω set corresponds to all pairs (e_{In}, λ) such that
i) for $\lambda \in (\lambda_{\min}, \lambda_{\max}]$

$$e_{In} \in \left[\frac{3}{\theta_3} - \sqrt{\left(\frac{3}{\theta_3} - \frac{N}{\bar{\theta}}\right)^2 - \frac{2\lambda}{\theta_3}}, \frac{3}{\theta_3} + \sqrt{\left(\frac{3}{\theta_3} - \frac{N}{\bar{\theta}}\right)^2 - \frac{2\lambda}{\theta_3}}\right],$$

ii) for $\lambda \in (0, \lambda_{\min}]$

$$e_{In} \in \begin{cases} \left[\frac{3}{\theta_3} - \sqrt{\left(\frac{3}{\theta_3} - \frac{N}{\bar{\theta}}\right)^2 - \frac{2\lambda}{\theta_3}}, \frac{N-1}{\theta_{N-1}} - \sqrt{\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\bar{\theta}}\right)^2 - \frac{2\lambda}{\theta_{N-1}}}\right) \\ \cup \left(\frac{N-1}{\theta_{N-1}} + \sqrt{\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\bar{\theta}}\right)^2 - \frac{2\lambda}{\theta_{N-1}}}, \frac{3}{\theta_3} + \sqrt{\left(\frac{3}{\theta_3} - \frac{N}{\bar{\theta}}\right)^2 - \frac{2\lambda}{\theta_3}}\right] \end{cases},$$

iii) for $\lambda = 0$,

$$e_{In} \in \left[2\frac{N-1}{\theta_{N-1}} - \frac{N}{\bar{\theta}}, 2\frac{3}{\theta_3} - \frac{N}{\bar{\theta}} \right]$$

where

$$\lambda_{\min} = \frac{\theta_{N-1}}{2}\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\bar{\theta}} \right)^2,$$

$$\lambda_{\max} = \frac{\theta_3}{2}\left(\frac{3}{\theta_3} - \frac{N}{\bar{\theta}} \right)^2.$$

A.5 Proof of Proposition 4

The participation constraint (5) of type θ_n, the marginal member in an inner club of size n, can be rewritten as

$$n\left(e_{In} - N/\bar{\theta}\right) - \lambda \geq \frac{\bar{\theta}(n-1) + \underline{\theta}(N-n)}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2$$

or, equivalently

$$n\left(e_{In} - N/\bar{\theta}\right)\left(1 - \frac{\bar{\theta}-\underline{\theta}}{N-1}\frac{\left(e_{In} + N/\bar{\theta}\right)}{2}\right) \geq \frac{N\underline{\theta} - \bar{\theta}}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2 + \lambda. \quad (9)$$

Similarly the non-participation constraint of the type θ_{n+1} can be written as

$$[n+1]\left(e_{In} - N/\bar{\theta}\right)\left(1 - \frac{\bar{\theta}-\underline{\theta}}{N-1}\frac{\left(e_{In} + N/\bar{\theta}\right)}{2}\right) < \frac{N\underline{\theta} - \bar{\theta}}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2 + \lambda. \quad (10)$$

(As before, if type $n+1$ is indifferent, she does not join). Define a function of x

$$F(x) = x\left(e_{In} - N/\bar{\theta}\right)\left(1 - \frac{\bar{\theta}-\underline{\theta}}{N-1}\frac{\left(e_{In} + N/\bar{\theta}\right)}{2}\right).$$

As the LHS of inequalities (9) and (10) coincide with $F(n)$ and $F(n+1)$ respectively, an inner club of size n can form in equilibrium if

$$F(x) \geq \frac{N\underline{\theta} - \bar{\theta}}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2 + \lambda \quad (11)$$

holds for $x = n$, but fails for $x = n+1$.

We begin by proving uniqueness. As by construction $e_{In} > N/\bar{\theta}$,

$$1 - \frac{\bar{\theta}-\underline{\theta}}{N-1}\frac{\left(e_{In} + N/\bar{\theta}\right)}{2} < 1 - \frac{\bar{\theta}-\underline{\theta}}{N-1}\left(\frac{N/\bar{\theta} + N/\bar{\theta}}{2}\right).$$

Given that the types are heterogeneous ($N < \bar{\theta}/\underline{\theta}$),

$$1 - \frac{\bar{\theta} - \underline{\theta}}{N-1}\left(\frac{N/\bar{\theta} + N/\bar{\theta}}{2}\right) = \frac{(N-1)\,\bar{\theta} - N\bar{\theta} + N\underline{\theta}}{N-1} = \frac{N\underline{\theta} - \bar{\theta}}{N-1} < 0.$$

Thus, the coefficient of x in $F(x)$ is negative, that is, $F(x)$ is decreasing in x. As the RHS of (11) is a constant for given model parameters and e_{In}, there will be at most one n such that

$$\begin{aligned} F(n) &\geq \frac{N\underline{\theta} - \bar{\theta}}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2 + \lambda, \\ F(n+1) &< \frac{N\underline{\theta} - \bar{\theta}}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2 + \lambda, \end{aligned} \tag{12}$$

which proves the uniqueness part.

To prove existence, we need to show that for any pair $(e_{In}, \lambda) \in \Omega$ the following two conditions hold:

$$F(3) \geq \frac{N\underline{\theta} - \bar{\theta}}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2 + \lambda, \tag{13}$$

$$F(N-1) < \frac{N\underline{\theta} - \bar{\theta}}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2 + \lambda, \tag{14}$$

From the continuity of $F(.)$ it follows that there exists a $n \in [3, N-2]$ such that the system (12) holds, which, in turn, implies that this n is the equilibrium size of the inner club.

We start by showing that inequality (13) holds for any $(e_{In}, \lambda) \in \Omega$. By definition of $F(.)$ inequality (13) can be rewritten as

$$3\left(e_{In} - N/\bar{\theta}\right) - \lambda \geq \theta_3\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2.$$

Solving for e_{In} shows that inequality (13) is satisfied for any (λ, e_{In}) such that

$$\lambda \leq \frac{\theta_3}{2}\left(\frac{3}{\theta_3} - \frac{N}{\overline{\theta}}\right)^2 \tag{15}$$

and

$$e_{In} \in \left[\frac{3}{\theta_3} - \sqrt{\left(\frac{3}{\theta_3} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_3}}, \frac{3}{\theta_3} + \sqrt{\left(\frac{3}{\theta_3} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_3}}\right]. \tag{16}$$

The set Ω is clearly a subset of the set defined by (15) and (16) as the set Ω is determined by further restrictions in addition to the conditions (15) and (16). Thus, as the inequality (13) holds for any pair (e_{In}, λ) satisfying (15) and (16), it also holds for any pair $(e_{In}, \lambda) \in \Omega$.

Now we show that condition (14) holds for any pair $(e_{In}, \lambda) \in \Omega$. Similarly to above, (14) is equivalent to

$$(N-1)\left(e_{In} - N/\bar{\theta}\right) - \lambda < \theta_{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2.$$

Solving for e_{In} yields that the inequality (14) is satisfies for any (λ, e_{In}) such that

$$\lambda > \frac{\theta_{N-1}}{2}\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}\right)^2 \tag{17}$$

and

$$e \in (-\infty, -\infty), \tag{18}$$

or for any pair (λ, e_{In}) such that

$$\lambda \leq \frac{\theta_{N-1}}{2}\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}\right)^2 \tag{19}$$

and

$$e_{In} \in \begin{cases} \left(-\infty, \frac{N-1}{\theta_{N-1}} - \sqrt{\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{N-1}}}\right) \\ \cup \left(\frac{N-1}{\theta_{N-1}} + \sqrt{\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{N-1}}}, \infty\right). \end{cases} \tag{20}$$

As above, the set Ω is a subset of the set determined by inequalities (17),(18), (19) and (20). Indeed, for both $\lambda > \frac{\theta_{N-1}}{2}\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}\right)^2$ and $\lambda \leq \frac{\theta_{N-1}}{2}\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}\right)^2$ there are additional restrictions imposed on e_{In} in the set Ω. Thus, inequality (14) is satisfies for any pair $(e_{In}, \lambda) \in \Omega$.

A.6 Characterization of the Ω_n Sets

We partition the set Ω into subsets Ω_n each corresponding to the pairs (λ, e_{In}) consistent with an inner organization of size n. Define

$$\lambda_n = \frac{\theta_n}{2}\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2.$$

For $n \in \{3, ..., N-2\}$ denote by Ω_n the set of all pairs (λ, e_{In}) such that
i) for $\lambda \in (\lambda_{n+1}, \lambda_n]$

$$e_{In} \in \left[\frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}, \frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}\right],$$

ii) for $\lambda \in (0, \lambda_{n+1}]$

$$e_{In} \in \begin{cases} \left[\frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}} \right)^2 - \frac{2\lambda}{\theta_n}}, \frac{n+1}{\theta_{n+1}} - \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}} \right)^2 - \frac{2\lambda}{\theta_{n+1}}} \right) \\ \cup \left(\frac{n+1}{\theta_{n+1}} + \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}} \right)^2 - \frac{2\lambda}{\theta_{n+1}}}, \frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}} \right)^2 - \frac{2\lambda}{\theta_n}} \right] \end{cases},$$

iii) for $\lambda = 0$

$$e_{In} \in \left[2\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}, 2\frac{n}{\theta_n} - \frac{N}{\overline{\theta}} \right].$$

Proposition A1: *Provided that types are heterogenous, an inner club of size n can form in equilibrium iff the pair (e_{In}, λ) belongs to the set Ω_n,* where $n \in \{3, ..., N-2\}$.

Proof. An inner organization of size $n \in [3, N-2]$ is a Nash equilibrium if the constraints (5) and (6) are satisfied and $e_{In} > N/\overline{\theta}$ holds. Solving inequality (5) we obtain

$$e_{In} \in \begin{cases} \emptyset & \text{for } \lambda > \lambda_n, \\ \left[\frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}} \right)^2 - \frac{2\lambda}{\theta_n}}, \frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}} \right)^2 - \frac{2\lambda}{\theta_n}} \right] & \text{for } \lambda_n \geq \lambda. \end{cases} \tag{21}$$

Similarly, inequality (6) yields

$$e_{In} \in \begin{cases} (-\infty, \infty) & \text{for } \lambda > \lambda_{n+1}, \\ \left(-\infty, \frac{n+1}{\theta_{n+1}} - \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}} \right)^2 - \frac{2\lambda}{\theta_{n+1}}} \right] & \\ \cup \left[\frac{n+1}{\theta_{n+1}} + \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}} \right)^2 - \frac{2\lambda}{\theta_{n+1}}}, \infty \right) & \text{for } \lambda_{n+1} \geq \lambda. \end{cases} \tag{22}$$

Before solving the system (21) and (22) and checking whether $e_{In} > N/\overline{\theta}$, we establish two useful results.

Lemma A2: $\lambda_n > \lambda_{n+1}$

Proof. Consider the function

$$\Lambda(x) = \frac{1}{2} \left(\frac{N-x}{N-1} \underline{\theta} + \frac{x-1}{N-1} \overline{\theta} \right) \left(\frac{x}{\left(\frac{N-x}{N-1} \underline{\theta} + \frac{x-1}{N-1} \overline{\theta} \right)} - \frac{N}{\overline{\theta}} \right)^2$$

and note that $\Lambda(n) = \lambda_n$ and $\Lambda(n+1) = \lambda_{n+1}$.

$$\frac{\partial \Lambda(x)}{\partial x} = -\frac{\left(\overline{\theta} - N\underline{\theta}\right)^2}{2\overline{\theta}^2} \frac{2\left(N-x\right)\left((N-x)\underline{\theta} + (N+x-2)\overline{\theta}\right)}{(N-1)\left(\frac{N-x}{N-1}\underline{\theta} + \frac{x-1}{N-1}\overline{\theta}\right)^2} < 0 \quad \text{if} \quad 2 \leq x < N$$

Thus $\lambda_n > \lambda_{n+1}$. ■

Lemma A3: *For every $\lambda \leq \lambda_{n+1}$, it holds that*

$$\frac{n+1}{\theta_{n+1}} - \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}} > \frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}} \tag{23}$$

and

$$\frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}} > \frac{n+1}{\theta_{n+1}} + \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}} \tag{24}$$

Proof. Consider first inequality (23). We define a function $F_1(\lambda)$ on $[0, \lambda_{n+1}]$ where

$$F_1(\lambda) = \frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}} - \left(\frac{n+1}{\theta_{n+1}} - \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}\right)$$

At $\lambda = 0$

$$F_1(0) = \frac{n}{\theta_n} - \left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right) - \left(\frac{n+1}{\theta_{n+1}} - \left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)\right) = 0$$

Moreover, whenever defined

$$\frac{\partial F_1(\lambda)}{\partial \lambda} = \frac{1}{\theta_n \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}} - \frac{1}{\theta_{n+1}\sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}} < 0$$

Indeed

$$\theta_n \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}} > \theta_{n+1}\sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}$$

as

$$\theta_n^2 \left(\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}\right) - \theta_{n+1}^2 \left(\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}\right)$$

$$= \left(\frac{(N-n)\left(\overline{\theta} - N\underline{\theta}\right)}{\overline{\theta}(N-1)}\right)^2 - \left(\frac{(N-n-1)\left(\overline{\theta} - N\underline{\theta}\right)}{\overline{\theta}(N-1)}\right)^2 - 2\lambda\left[\theta_n - \theta_{n+1}\right]$$

$$= \left(\frac{\left(\overline{\theta} - N\underline{\theta}\right)}{\overline{\theta}(N-1)}\right)^2 (2N - 2n - 1) + 2\lambda \frac{\left(\overline{\theta} - \underline{\theta}\right)}{(N-1)} > 0$$

Thus, $F_1(\lambda)$ is a decreasing function of λ and $F_1(\lambda)|_{\lambda>0} < 0$. This proves inequality (23).

Similarly, to prove inequality (24) we define a function $F_2(\lambda)$ on $[0, \lambda_{n+1}]$ where

$$F_2(\lambda) = \frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}} - \left(\frac{n+1}{\theta_{n+1}} + \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}\right)$$

From Lemma A1 it follows that at $\lambda = 0$

$$F_2(0) = 2\left(\frac{n}{\theta_n} - \frac{n+1}{\theta_{n+1}}\right) > 0$$

Moreover,

$$\frac{\partial F_2(\lambda)}{\partial \lambda} = -\frac{1}{\theta_n\sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}} + \frac{1}{\theta_{n+1}\sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}} = -\frac{\partial F_1(\lambda)}{\partial \lambda} > 0$$

Thus, $F_2(\lambda)$ is a increasing function of λ and $F_2(\lambda)|_{\lambda>0} > 0$. This is equivalent to

$$\frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}} > \frac{n+1}{\theta_{n+1}} + \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}},$$

thereby proving inequality (24). ■

Using Lemma A2 and A3, we can now describe the entire set of joint solutions for (21) and (22). For $\lambda > \lambda_n$, the inequality (21) and hence the system has no solution.

For $\lambda_n \geq \lambda > \lambda_{n+1}$, the intersection of (21) and (22) results in

$$e_{In} \in \left[\frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}, \frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}\right] \tag{25}$$

According to Lemma A3,

$$\frac{n+1}{\theta_{n+1}} - \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}} > \frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}$$

and

$$\frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}} > \frac{n+1}{\theta_{n+1}} + \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}.$$

Hence, the intersection of (21) and (22) for $\lambda_{n+1} \geq \lambda > 0$ is

$$e_{In} \in \begin{cases} \left[\frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}, \frac{n+1}{\theta_{n+1}} - \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}\right) \\ \cup \left(\frac{n+1}{\theta_{n+1}} + \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}, \frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}\right] \end{cases} \tag{26}$$

For each $\lambda > 0$

$$\frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}} > \frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2} = \frac{N}{\overline{\theta}}.$$

Thus, $e_{In} > N/\overline{\theta}$ is satisfied for any e_{In} belonging to the sets (25) and (26).

Imposing the restriction $e_{In} > N/\overline{\theta}$ on the intersection of (21) and (22) for $\lambda = 0$ gives

$$e_{In} \in \left[2\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}, 2\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right]. \tag{27}$$

Thus, (25), (26) and (27) describe the set Ω_n. This concludes the proof of Proposition A1.

A.7 Proof of Proposition 6

Proof by contradiction. If an equilibrium with an inner organization of size $n < N$ exists,

$$n(e_{In} - e_{Out}) - \lambda > \theta_n \left[e_{In}^2 - e_{Out}^2\right]/2 \tag{28}$$

and

$$e_{In} > e_{Out}$$

must hold where

$$e_{Out} = \frac{N}{m\overline{\theta} + (1-m)\underline{\theta}}$$

is the best preferred outer-club effort of the decisive member under the majority rule $\mathcal{M}(m)$. Setting $e_{In} = N/(m\overline{\theta} + (1-m)\underline{\theta}) + \delta$ and inserting it in equation (28) yields

$$n\delta - \lambda \geq \frac{\theta_n}{2}\left(2\delta\frac{N}{m\overline{\theta} + (1-m)\underline{\theta}} + \delta^2\right),$$

or equivalently,

$$\delta\left(\frac{n}{\theta_n} - \frac{N}{m\overline{\theta} + (1-m)\underline{\theta}}\right) \geq \frac{\delta^2}{2} + \frac{\lambda}{\theta_n} > 0.$$

This condition can only be satisfied if

$$\frac{n}{\theta_n} > \frac{N}{m\overline{\theta} + (1-m)\underline{\theta}}. \tag{29}$$

Inserting the explicit expression (1) for θ_n and rearranging yields

$$m > 1 - \frac{1}{n}\frac{(N-n)\left(\overline{\theta} - N\underline{\theta}\right)}{(N-1)\left(\overline{\theta} - \underline{\theta}\right)}. \tag{30}$$

As

$$\frac{(N-n)\left(\overline{\theta} - N\underline{\theta}\right)}{(N-1)\left(\overline{\theta} - \underline{\theta}\right)} < 1, \tag{31}$$

inequality (30) implies that the necessary condition for formation of an inner club of size n is

$$m > 1 - \frac{1}{n} \geq \frac{1}{2}, \tag{32}$$

as an inner club should have at least two members. Simple majority threshold $m = 1/2$ never satisfies condition (32). We conclude that under the majority threshold $m = 1/2$ any divided organization ceases to exist.

A.8 Proof of Proposition 7

Consider condition (30). The RHS of it is increasing in n (as $N < \overline{\theta}/\underline{\theta}$). Therefore, if condition (30) fails for $n = 2$, i.e. club of size 2 ceases to exist, then any larger club also ceases to exist. Therefore no divided organization emerges as long as the majority threshold m prevents formation of divided organization of size 2. Now denote

$$\overline{m}(\underline{\theta}, \overline{\theta}, N) = 1 - \frac{(N-2)\left(\overline{\theta} - N\underline{\theta}\right)}{2\,(N-1)\left(\overline{\theta} - \underline{\theta}\right)}.$$

By condition (30) and discussion above any majority rule $\mathcal{M}(m)$ with $m < \overline{m}(\underline{\theta}, \overline{\theta}, N)$ results in no divided organization forming in equilibrium. Further, using inequality (31) one can see that

$$\overline{m}(\underline{\theta}, \overline{\theta}, N) = 1 - \frac{(N-2)\left(\overline{\theta} - N\underline{\theta}\right)}{2\,(N-1)\left(\overline{\theta} - \underline{\theta}\right)} > 1 - \frac{1}{2} = 1/2.$$

Finally, consider an increase in the agents' heterogeneity via a change in the support of the distribution of types. Then higher heterogeneity (higher $\overline{\theta}$ and lower $\underline{\theta}$) corresponds to a lower $\overline{m}(\underline{\theta}, \overline{\theta}, N)$

$$\partial \frac{\overline{m}(\underline{\theta}, \overline{\theta}, N)}{\partial \underline{\theta}} > 0, \qquad \partial \frac{\overline{m}(\underline{\theta}, \overline{\theta}, N)}{\partial \overline{\theta}} < 0.$$