Optimal Voting Rules for International Organizations, with an
Application to the UN

Johann Caro Burnett

November 24, 2016

Abstract

This paper examines a self-enforcing mechanism for an international organization that interacts repeatedly over time. A random shock determines which countries would be in favor of or against taking a collective action. If the organization wants to take the action, some countries may disagree with participating; and therefore, incentives must be provided to such countries. I study a range of equilibria, varying from arbitrarily patient to myopically impatient members. Moreover, in order to find a simple closed form solution to the equilibrium, I restrict attention to stationary payoffs. I characterize the optimal mechanism within this class of equilibria. Then I show that this mechanism can be implemented by a weighted voting rule. Finally, I contrast the findings with the United Nations’ voting system.
1 Introduction

International Organizations have an increasingly important role in the global decision making. The decision-making-rule of each of these organizations will determine the outcomes that will (arguably) affect several countries, if not the entire world. This paper studies such decision making rules. From an ideal economic point of view, we want this decision making rule to be a mechanism that implements efficient actions. However, the voting rules we observe in these International Organizations are very different from each other. For instance, The UN Security Council has five permanent members who hold veto power, and ten rotating members. The World Bank, International Monetary Fund, European Union have a weighted vote system. The WTO and the NATO use Unanimity. Two natural questions that arise are from the diversity in these voting rules are, what is the best way to make collective decisions in an International Organization? What set of rules implement the best possible outcome, and at the same time provide incentives to the countries involved in order for them to participate?

I will study a stochastic game with three elements that International Organizations may typically have. First, they cannot rely on external enforcers. Thus, any set of rules they use must be self-enforcing. Second, members are heterogeneous; countries have large differences in income, military power, natural resources, etc. Some countries have a stronger opinion on global problems and other countries care mostly about their local issues. Third, the organization cannot use/rely on monetary

\(^1\)They work on a variety of global issues such as preserving peace (United Nations), environment (The two most relevant international agreements, yet not precisely organizations, are the Montreal Protocol and Kyoto Protocol), migration (International Organization for Migration, International Centre for Migration Policy Development), weapons control (Organization for the Prohibition of Chemical Weapons, International Atomic Energy Agency), law enforcement (International Criminal Court), and many other issues. Some of these organizations are arguably very influential. For instance, one of the branches of the United Nations (UN), the Security Council, can authorize military actions such as the Gulf War (See the Security Council’s resolution 678), impose sanctions such the recent ones on North Korea and Iran over their nuclear programs(See the Security Council’s resolution 1737 and 1874), and some other actions such as the no-fly zone over Libya(See the Security Council’s resolution 1973).

\(^2\)In a stochastic game, the players interact repeatedly over time. But different from a repeated game, there is a stochastic shock that affects the payoff of the players.
transfers.\footnote{This third assumption may seem the most restrictive in this paper. However, there are many reasons to justify the absence of transfers. First, transfers are, in general, not openly used (if used at all). For example, the UN charter does not mention monetary transfers between countries as a way of compensation towards affected countries. There are some studies (For example, see Kuziemko and Werker (2006)) showing that being elected as a non permanent member in the UN Security Council is correlated with foreign aid. However, there is no evidence of causality. Moreover, foreign aid usually imposes several restrictions. For instance, the resources may be targeted (towards health, education, etc.), or there could be implicit inefficiencies (bureaucracy, corruption, etc.). Additionally, transfers do not necessarily solve the provision of incentives in a trivial way. Any transfer has to be self enforcing itself, so countries have to be willing to complain with any transfer schedule as well. This may introduce additional constraints and therefore go beyond the scope of the present study.} Note that, neither the absence of transfers nor the perfectly and unrestricted use of them are realistic assumptions; and in practice we would be somewhere in between those two cases. However, I want to examine the provision of incentives purely by choosing the appropriate decision rule.

\textbf{Literature}

This paper is mostly related to the existing literature on endogenous decision rules. Maggi and Morelli (2006) focus on efficiency and self-enforceability. The difference from their paper is that here the members of the organization are heterogeneous and I allow for non stationary equilibria. With these two extensions, I can provide one explanation why some organizations use different weights for their members (IMF, World Bank), as well as why some organizations have some sort of randomness in their decision power (UN Security Council). Barbera and Jackson (2004) study stability of decision rules. In their paper, a decision rule is stable if it would choose itself when voted against other decision rules. In this paper, the goal of the organization is not to generate stability but to maximize the sum of members’ payoffs. While maximizing the sum of payoffs, I also indirectly endogenize the decision rule. However, the criteria for choosing one particular decision rule here is different from them. While in their paper the self-stable voting rule is simple majority (or something very near simple majority), in this paper the voting rule is state dependent, each country has different weights, and the threshold for implementing an action is not necessarily 50\% of the votes. 
Voeten (2001) describes the bargaining power of the members of the Security Council as a function of outside options. He shows how a country can achieve better outcomes with higher outside options. Although his model does not study Security Council elections, it says that the voting power increases with higher outside options. This work goes in line with one of the extensions in the model. Namely, when I allow for heterogeneous outside options, these outside options have a positive effect on a country’s voting power. Dreher et al. (2014) study the determinants of the election at the Security Council. They show that GNP and population, as well as the number of years being outside the Security Council have, have a positive effect on the probability of being elected as a non-permanent member. Their paper is very close to the present work in two ways. First, country characteristics can be seen as proxies to outside option. Second, their ‘turn-taking’ variable says that the longer a country is not elected, the more likely it gets to being a member of the Security Council. We can rationalize such result within the framework of the model studied here. A member who has currently no voting power knows that the decisions made at the organization are poorly correlated to its preferences; and therefore unfavorable actions are taken with a high probability. This means that with a high probability, the organization has to promise a higher voting weight to such country in order to secure its participation.

**Motivation**

In reality, we observe a wide range of voting rules in International Organizations. For example, the IMF, World Bank, and European Union use voting weights that are (almost) constant through time and depend on specific variables (i.e. contributions to the organization). On the other hand, the NATO and WTO follow a unanimity rule. Moreover, in all the previous cases, each country knows
how much its vote worths and that will not change from period to period.

In contrast, the United Nations follows a completely different voting rule.\textsuperscript{4} Each year, only a subset of the members, called the Security Council, votes on relevant and compulsory issues. Moreover, except from the five permanent members\textsuperscript{5} who are always part of the Security Council, there is uncertainty on which countries will have the right to vote in the Security in the future. Historically, however, some countries have been part of the Security Council more often than others. The table below shows a list of the countries who have been elected more often:

Table 1: Top 19 countries most often in the Security Council

<table>
<thead>
<tr>
<th>Rank</th>
<th>Region</th>
<th>Country</th>
<th>Elected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Asia and Pacific</td>
<td>Japan</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>Latin America</td>
<td>Brazil</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Latin America</td>
<td>Argentina</td>
<td>9</td>
</tr>
<tr>
<td>4-7</td>
<td>Latin America</td>
<td>Colombia</td>
<td>7</td>
</tr>
<tr>
<td>4-7</td>
<td>Asia and Pacific</td>
<td>India</td>
<td>7</td>
</tr>
<tr>
<td>4-7</td>
<td>Western Europe</td>
<td>Italy</td>
<td>7</td>
</tr>
<tr>
<td>4-7</td>
<td>Asia and Pacific</td>
<td>Pakistan</td>
<td>7</td>
</tr>
<tr>
<td>8-10</td>
<td>Western Europe</td>
<td>Canada</td>
<td>6</td>
</tr>
<tr>
<td>8-10</td>
<td>Africa</td>
<td>Egypt</td>
<td>6</td>
</tr>
<tr>
<td>8-10</td>
<td>Western Europe</td>
<td>Germany</td>
<td>6</td>
</tr>
<tr>
<td>11-19</td>
<td>Western Europe</td>
<td>Australia</td>
<td>5</td>
</tr>
<tr>
<td>11-19</td>
<td>Western Europe</td>
<td>Belgium</td>
<td>5</td>
</tr>
<tr>
<td>11-19</td>
<td>Latin America</td>
<td>Chile</td>
<td>5</td>
</tr>
<tr>
<td>11-19</td>
<td>Western Europe</td>
<td>Netherlands</td>
<td>5</td>
</tr>
<tr>
<td>11-19</td>
<td>Africa</td>
<td>Nigeria</td>
<td>5</td>
</tr>
<tr>
<td>11-19</td>
<td>Latin America</td>
<td>Panama</td>
<td>5</td>
</tr>
<tr>
<td>11-19</td>
<td>Eastern Europe</td>
<td>Poland</td>
<td>5</td>
</tr>
<tr>
<td>11-19</td>
<td>Western Europe</td>
<td>Spain</td>
<td>5</td>
</tr>
<tr>
<td>11-19</td>
<td>Latin America</td>
<td>Venezuela</td>
<td>5</td>
</tr>
</tbody>
</table>

The five permanent members were excluded from the ranking. Some non-European countries have been included in the Western European group, this includes Canada, Australia, and Israel. The classification of regions has been changed once in 1966. I use the current classification for simplicity.

\textsuperscript{4}For a more detailed explanation on the UN, see section 3
\textsuperscript{5}China, France, Russia, United Kingdom, and United States.
Moreover, regions vary in their number of countries as well as their fixed\(^6\) number of seats. For example, Germany has been elected six times, and Poland five times. However, Germany belongs to the Western European region which has two seats at the Security Council, and Poland belongs to the Eastern European region which has only one seat at the Security Council. Therefore, one can see that conditional on region, Poland seems to be more likely to get elected than Germany. To see how the regions change the perspective, table 2 accounts for regional disparity.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Region</th>
<th>Country</th>
<th>Prob. of election cond. on region</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Eastern Europe</td>
<td>Poland</td>
<td>0.178</td>
</tr>
<tr>
<td>2</td>
<td>Asian and Pacific</td>
<td>Japan</td>
<td>0.166</td>
</tr>
<tr>
<td>3</td>
<td>Eastern Europe</td>
<td>Romania</td>
<td>0.142</td>
</tr>
<tr>
<td>3</td>
<td>Eastern Europe</td>
<td>Ukraine</td>
<td>0.142</td>
</tr>
<tr>
<td>5</td>
<td>Latin America</td>
<td>Brazil</td>
<td>0.136</td>
</tr>
<tr>
<td>6</td>
<td>Latin America</td>
<td>Argentina</td>
<td>0.123</td>
</tr>
<tr>
<td>7</td>
<td>Eastern Europe</td>
<td>Bulgaria</td>
<td>0.107</td>
</tr>
<tr>
<td>7</td>
<td>Eastern Europe</td>
<td>Czech Republic</td>
<td>0.107</td>
</tr>
<tr>
<td>9</td>
<td>Asia and Pacific</td>
<td>India</td>
<td>0.106</td>
</tr>
<tr>
<td>9</td>
<td>Asia and Pacific</td>
<td>Pakistan</td>
<td>0.106</td>
</tr>
<tr>
<td>11</td>
<td>Latin America</td>
<td>Colombia</td>
<td>0.095</td>
</tr>
<tr>
<td>12</td>
<td>Western Europe</td>
<td>Italy</td>
<td>0.093</td>
</tr>
<tr>
<td>13-14</td>
<td>Western Europe</td>
<td>Canada</td>
<td>0.080</td>
</tr>
<tr>
<td>13-14</td>
<td>Western Europe</td>
<td>Germany</td>
<td>0.080</td>
</tr>
<tr>
<td>15-16</td>
<td>Africa</td>
<td>Egypt</td>
<td>0.071</td>
</tr>
<tr>
<td>15-16</td>
<td>Eastern Europe</td>
<td>Hungary</td>
<td>0.071</td>
</tr>
</tbody>
</table>

The rank is worldwide, and based on the probability displayed in column four. The probability of election conditional on region is the ratio of total number of elections of a country divided by the total number of elections of the region that the country belongs to. The five permanent members of the Security Council as well as Former UN members were excluded. The classification of regions has been changed once in 1966. I use the current classification for simplicity.

From these two tables, we can see two things. First, voting power in the Security Council is heterogeneous even if we only look at non permanent members. Some countries are part of the

\(^6\)Three for Africa, two for Asia-Pacific, two for Latin America, two for West Europe, and one for East Europe.
Security Council much more often than others. Indeed, out of the 193 current UN members, 73 have never been part of the Security Council. Second, there is randomness in the voting power. A country can have a lot of power during a few years, and then have basically zero voting power for several years afterwards. A more specific research question from observing the United Nations would be whether we can rationalize this randomness as an optimal equilibrium.

The rest of the paper is divided as follows: First, on section 2, I describe and solve the model. As a benchmark, I will characterize an efficient outcome assuming that actions are enforceable. Then I show that, this efficient outcome can be implemented by a weighted voting rule. This Pareto efficient allocation will be useful to compare with the solution of the optimal mechanism for the stochastic game. In order to find a simple closed form solution to the equilibrium, I restrict attention to stationary payoffs. The solution to the optimal mechanism matches the Pareto efficient allocation whenever the members of the organization are sufficiently patient. Moreover, regardless of whether they are patient or impatient, the mechanism can be implemented by a weighted voting rule. Then, I show how these voting rules are quite general by mapping them into most of the known voting rules such as simple majority rule, oligarchy, and dictatorship; and they can also have other usual properties such as veto power and unanimity. Finally, I will discuss some extensions to the model. In section 3 I briefly describe the UN voting system, and contrast it with the findings of the theoretical model. Finally, I conclude in section 4.
2 The Model

2.1 The Stage Game

There are $N$ countries endowed with a binary action space; they can choose to either participate or not on a (pure) collective action. That is, if everyone participates, the collective action is effective. Conversely, if at least one of the countries decides not to participate, the action fails and the status quo is preserved. At the beginning of the period, the state of the world realizes. This state of the world will be denoted as $y = (y_1, y_2, \ldots, y_N)$, and is the profile of payoffs of all members in the case the collective action is taken. That is, when the collective action is effective, each member receives a payoff $y_i$, which is iid across countries and periods. If the action is not taken, everybody gets their status quo payoff, which is normalized to zero. Country $i$’s payoff ($y_i$) can take one of two values. With probability $p$, it takes a high value $y_i > 0$, and with probability $1 - p$, it takes a low value $y_i < 0$.

A country is in favor (against) taking the action whenever its payoff is higher (lower) than the status quo payoff.

2.2 Perfect Enforceability Benchmark

As a benchmark, let us consider all the Pareto efficient allocations. Given a profile of Pareto weights $(\lambda_i)$, we will characterize the best outcome assuming that the action are enforceable. At any state of the world $y$, the Pareto efficient allocation is the solution to the following problem:

$$\max_{x \in \{0,1\}} \sum_{i=1}^{N} \lambda_i \tilde{y}_i(y)$$

where $y_i(y)$ is the preference shock of the $i^{th}$ country on state $y$. Clearly, it is optimal to take the
collective action whenever the sum in the expression above is positive, and to preserve the status quo when the sum above is negative.

**Lemma 1.** Let \(- \sum \lambda_i y_i\) be the worst possible (aggregated) loss from taking the collective action, and let \(\bar{y}_i - y_i\) be country \(i\)'s gain from taking the collective action in a state that favors such country. Then the Pareto efficient rule is to take the action if the weighted sum of the gains of all countries favored in the current state exceeds the worst possible loss. That is, whenever the following condition holds:

\[
\sum_{i \mid y_i(y) > 0} \lambda_i(\bar{y}_i - y_i) \geq -\sum_{i=1}^{N} \lambda_i y_i
\]

**Proof.** Let \(a_i(y)\) indicate whether country \(i\) had a positive preference shock. Then, each country's preference shock can be expressed as \(y_i(y) = a_i(y)\bar{y}_i + (1 - a_i(y))y_i = a_i(y)(\bar{y}_i - y_i) + y_i\). With this new notation, I can rewrite the Pareto efficient maximization problem in the following way:

\[
\max_{x \in \{0,1\}} x \sum_{i=1}^{N} \lambda_i \left( a_i(y)(\bar{y}_i - y_i) + y_i \right)
\]

It is optimal to take the action whenever the expression above is positive, that is, whenever \(\sum_{i=1}^{N} \lambda_i a_i(y)(\bar{y}_i - y_i) \geq -\sum_{i=1}^{N} y_i\). Then the result follows, as \(a_i(y) = 1\) on states such that \(y_i(y) > 0\) and zero otherwise.

Finally, let us make a remark on the Pareto frontier. If the decision rule is binary, the Pareto frontier consists of a finite set of points. If we were to allow the decision variable to take values on the \([0,1]\) interval, the Pareto frontier would be convex. In either case, small perturbations in the Pareto weights do not change, in general, the Pareto-optimal decision rule. This can be shown in the following...
Example 1

There are two countries \( \{A, B\} \) with ex-ante identical preferences: \( y_i = 2 \) and \( y_i = -1 \), and Lagrange multipliers such that \( \lambda_A + \lambda_B = 1 \). The worst possible loss is \(-\sum \lambda_i y_i = 1\). Let us start with one extreme case \( \lambda_A = 1 \). Here, it is clear that \( A \) is a dictator. In particular, the action is not taken in the state \((y_1, y_2) = (-1, 2)\). However, under an egalitarian decision rule, the action would be implemented in that state.

The next step in this example, would be to compute the smallest \( \lambda_A \) such that \( A \) is still a dictator. I solve this by making the maximization problem indifferent between taking the action or preserving the status quo in \((y_1, y_2) = (-1, 2)\). That happens when: \( \sum_{y_i(y) > 0} \lambda_i(y_i - y) = -\sum \lambda_i y_i \), or replacing the values \( \lambda_B(2 + 1) = 1 \), and therefore \( \lambda_B = 1/3 \) or \( \lambda_A = 2/3 \). For any \( \lambda_A = 2/3 - \epsilon \), the Pareto efficient allocation will implement the action in state \((y_1, y_2) = (-1, 2)\) as well as states \((2, -1)\) and \((2, 2)\). This is indeed the Egalitarian outcome, which is implemented not only for \( \lambda_A = 1/2 \), but for any \( \lambda_A \) in the range: \([1/3, 2/3]\).

Finally, by symmetry, it is easy to see that \( \lambda_A < 1/3 \) makes \( B \) a dictator. Thus, the three relevant voting rules in the example are A-dictatorship, B-dictatorship, and the egalitarian allocation. If we restrict the decision making rule to be discrete, there are only three efficient outcomes. However, we can obtain a convex Pareto frontier by convex combinations of A-dictatorship with egalitarian, and B-dictatorship with egalitarian. This is illustrated in figure 1.
Let us look at the payoff profile \((\tilde{u}_A, \tilde{u}_B)\). This payoff cannot be obtained by a discrete mechanism. However, it is still efficient when \(\lambda_A = 2/3\). This payoff can be obtaining by tossing a coin and with probability \(1/2\) make \(A\) a dictator, and with complimentary probability \(1/2\) implement the Egalitarian outcome. Alternatively, this outcome can also be implemented by setting \(x = 1\) on states \((2, 2)\) and \((2, -1)\); \(x = 1/2\) on state \((-1, 2)\); and \(x = 0\) on state \((-1, -1)\). The interpretation of \(x = 1/2\) is the following: in some states, the organization decides to compromise by implementing the action only partially.

### 2.3 Weighted Voting Rule

Before I describe the equilibrium of the game, let us propose an alternative and less abstract way to look at the Pareto efficient decision rule. To do this, first let us define a weighted voting rule as a profile of weights \(m\) and a target \(M\) such that: every country has a weight \(m_i\), countries vote on whether
they want to take the collective action, and the action is implemented if the sum of the weights of all members who voted in favor of taking the action exceeds a target $M$. Else, the outcome will be the status quo.

**Lemma 2.** For a given profile of Pareto weights, the efficient outcome can be implemented by a weighted voting rule.

*Proof.* I prove this lemma by constructing a profile of weights and a target that reflect the Pareto efficient. The construction follows in a very straightforward way from the characterization of the Pareto efficient. Namely, we set $m_i = \lambda_i(\bar{y}_i - n_i)$ and $M = -\sum \lambda_i y_i$. □

This lemma tells us that any efficient decision rule can be achieved by weighted votes. Moreover, the class of outcomes that can be implemented by weighted votes is fairly large. Indeed, weighted votes can include some well known examples such as the egalitarian decision rule, dictatorship, oligarchy, veto power, and one-country-one-vote. The following example illustrates this point:

**Example 2**

There are three countries $\{A, B, C\}$ with ex-ante identical preferences: $\bar{y}_i = 2$ and $y_i = -3$. Moreover, let us normalize the Pareto weights so that their sum equals one. To illustrate one simple case, take $\lambda_A = 1$. Then the Pareto efficient allocation will be to take the action if and only if country $A$ votes *yes*. Moreover, as I noted above, small perturbations in the Pareto weights do not change (in general) the decision rule. Thus, for $\lambda_A$ near to one, country $A$ will still be a dictator. Figure 2 shows all possible combinations of decision rules in this example. On the region labeled as $A-B$ Oligarchy, both $A$ and $B$ have veto power, and country $C$ is never pivotal. And, on the region labeled as $A$ Veto Power, country $A$ has enough weight to veto unfavorable decisions, but not enough weight to be a
dictator. Moreover in this region $B$ and $C$ can be pivotal. They change the outcome in the event that $A$ voted *yes* and the two other countries disagree with each other.

![Figure 2: Decision rules for all combinations of Pareto weights.](image)

2.4 The Repeated Game

The $N$ member of the International Organization interact repeatedly over time and discount time using a constant factor $\delta$. We will regard the organization as a mechanism that collects preferences and suggests an outcome. Therefore we will use the words ‘organization’ and ‘mechanism’ indistinctly. From the payoff structure, the International Organization has two alternatives on each period: they can either take the collective action or preserve the status quo. However, the organization cannot force the members to participate in its decision. That is, after observing the state of the world, each member individually decides whether to participate with the decision made by the organization.

Let us denote by $x$ as the indicator function for the recommended collective action, that is $x = 1$ if

---

7 Asking a subset of members to take the action has the same output as asking everyone not to participate.
the organization decides that all members should participate, and \( x = 0 \) otherwise. A member of the organization will receive a payoff equal to the present discounted sum of the streams of all its payoffs:

\[
(1 - \delta) \sum_{t=0}^{\infty} \delta^t x_t y_{i,t}
\]

Moreover, notice that the status quo payoff comes from not taking the action in the current period but staying in the organization. For simplicity, I assume that receiving the status quo payoff forever is the same as not having the organization at all.\(^8\) Thus I also set the outside option payoff equal to zero. Finally, I also assume that \( p \bar{y}_i + (1 - p) \underline{y}_i > 0 \). That is, the status quo payoff is smaller than the payoff of always taking the action. These two assumptions avoid corner solutions, but are not essential for the results of the model. Indeed, I will discuss the effects of explicit outside options on section 2.8.2.

### 2.5 The Self-Enforcing Optimal Voting Mechanism

I am interested in characterizing the optimal self enforcing decision rule within the class of stationary equilibria. The equilibrium concept I will use is Perfect Public Equilibrium, where the observable outcome variable is the profile of individual actions. There are three constraints to be satisfied. First, all countries must be willing to join and maintain their membership at the organization. Second, all members must report their preferences truthfully. Third, all members must be willing to participate by taking the action whenever the organization decides to do so. Next, I write the maximization problem and explain the restrictions:

\(^8\)Although I could make the outside option different from the status quo payoff.
\[
\max_x (1 - \delta) \sum_{t=0}^{\infty} \delta^t \sum_{y_t \in Y} Pr(y_t) x_t(y_t) \sum_{i \in N} \lambda_i y_{i,t} \\
\text{subject to}
\sum_{t=\tau+1}^{\infty} \delta^{t-\tau} \sum_{y_t \in Y} Pr(y_t) x_t(y_t) y_{i,t} \geq 0, \forall i
\]

Equation (2) is the voluntary membership constraint. It tells that after every history, the expected payoff of each member must be more desirable than leaving the organization, which we assumed yields zero payoff forever. Equation (3) is the truthtelling condition: members should report their true preferences. We should notice that for this constraint, the only way to provide incentives is by the payoffs of the current period. The reason is that we are restricting attention to stationary payoffs. Therefore, regardless of the current shock, the future payoff will be the same. Equation (4) is the participation constraint. It states that after every decision made by the organization, the members must be willing to participate in the organization’s decision. If the members comply, they receive an
instant payoff and a continuation payoff. The sum of these two payoffs must exceed the alternative, which is not to participate in the action and therefore receive the status quo payoff forever.\footnote{The instant payoff they receive from not taking the action is zero, and the continuation payoff is the status quo forever, as the organization breaks up.}

**Lemma 3.** The voluntary membership constraint (2) is not binding at the optimum.

**Proof.** If at the optimum, a member has veto power then the action is only implemented in states in which that country has a positive payoff (because that member is reporting truthfully at the optimum). Therefore the expected payoff of that member is a weighted sum of positive numbers and zeros. So it cannot be negative. Moreover, since the payoffs are stationary, this means that all future expected payoffs are the same and positive.

If a member is not a dictator, then there is at least one state where the action is implemented and the member has a negative payoff. Therefore, from equation (4), the future expected payoff has to be positive. Moreover, since the payoffs are stationary, this means that all future expected payoffs are the same and positive. Similar to the previous case, this has to hold for every expected payoff after every history.

\[\square\]

**Lemma 4.** \(x_t(y_{-i,t}, y_{i,t})\) weakly increasing in \(y_{i,t}\) is a sufficient condition to satisfy the truth-telling conditions in equation (3).

**Proof.** If \(x_t(y_{-i,t}, y_{i,t})\) is weakly increasing, then \(\sum_{y_{-i,t} \in Y_{-i}} Pr(y_{-i,t}) x_t(y_{-i,t}, y_{i,t})\) is also weakly increasing in \(y_{i,t}\). If \(y_{i,t} = \bar{y}_{i,t}\), then we need to show that:

\[
\sum_{y_{-i,t} \in Y_{-i}} Pr(y_{-i,t}) \left[ x_t(y_{-i,t}, \bar{y}_{i,t}) - x_t(y_{-i,t}, \bar{y}_{i,t}) \right] \bar{y}_{i,t} \geq 0
\]

But this follows immediately, as \(\sum_{y_{-i,t} \in Y_{-i}} Pr(y_{-i,t}) x_t(y_{-i,t}, y_{i,t})\) is weakly increasing.
Similarly, if \( y_{i,t} = y_{i,t} \), then we need to show that:

\[
\sum_{y_{-i,t} \in Y_{-i}} Pr(y_{-i,t}) \left[ x_t(y_{-i,t}, y_{i,t}) - x_t(y_{-i,t}, \overline{y_{i,t}}) \right] y_{i,t} \geq 0
\]

And again, this follows immediately, as \( \sum_{y_{-i,t} \in Y_{-i}} Pr(y_{-i,t}) x_t(y_{-i,t}, y_{i,t}) \) is weakly increasing, and \( y_{i,t} \) is negative.

\[\Box\]

### 2.6 Equilibrium

The first thing to notice is that there is always a solution to the problem. The payoffs are finite, the strategies are finite, and as studied in Maggi and Morelli (2006), the set of implementable outcomes is non empty. The non emptiness follows from the fact that the \textit{unanimity rule} is always feasible. To see this, we first note that it satisfies voluntary membership, as this voting rule implements the status quo in almost all the states (giving zero payoff to everyone) and it only implements the action whenever everyone agrees, which means everyone gets a small but positive expected payoff. Second, it satisfies the truth-telling condition, because the decision making rule is weakly increasing. Finally, unanimity also satisfies the participation constraint, because it only implements the action when everyone agrees. This means that even if the discount factor is zero, unanimity is still a solution, and it shall be regarded as another benchmark.

As mentioned above, a stationary equilibrium has the property that the expected payoffs do not change over time. A natural candidate for this equilibrium is the Pareto efficient allocation, together with grim trigger strategies. Notice that grim trigger strategies are not an assumption. Indeed, the best way to provide incentives is by punishing ‘off the equilibrium path’ behavior in the most severe
yet credible way. Namely, after observing a deviation from the equilibrium path, the organization
breaks up and all members receive the status quo payoff forever.

**Proposition 1.** There is a threshold $\delta^*$ such that if the discount factor exceeds that threshold, the
Pareto efficient allocation is the optimal mechanisms.

**Proof.** Notice that we only need to provide incentives to those countries who disagreed in taking the
action. Let us write the incentive constraint:

$$
(1 - \delta)y_i + \delta U_i^* \geq 0
$$

where $U_i^*$ is the Pareto Efficient expected payoff, which is positive as we assumed $p\bar{y}_i + (1 - p)y_i > 0$.
Moreover, the Pareto Efficient allocation also satisfies truth-telling, as reporting a high shock ($\bar{y}_i$)
only increases the probability of implementing the action. For each country that is not dictator, we
define: $\delta_i = -\frac{y_i}{y_i + U_i^*}$. Finally, the desired discount factor will be the largest of each country’s minimal
requirements: $\delta^* = \max\{\delta_1, \delta_2, \ldots, \delta_N\}$.

As a remark, note that if the Pareto efficient allocation is the solution to the problem restricted to
stationary payoffs, it is also the solution to the unrestricted problem, i.e. expected payoffs are allowed
to change over time. Next we study the solution to the problem when the discount factor is moderately
high. That is, it is smaller than $\delta^*$, but large enough so that it is still possible to implement an outcome
better than unanimity. The following lemma characterizes the decision making rule:

**Lemma 5.** Let $\gamma_i(y)$ be the Lagrange multiplier for the participation constrain (4), $\phi_i$ the Lagrange
multiplier for the truth-telling condition (3), and $\tilde{\phi}_i(y_i)$ be defined as follows:
\[ \tilde{\phi}_i(y_i) = \begin{cases} \phi_i & \text{if } y_i = \bar{y}_i \\ -\phi_i \frac{\gamma_i(y_i)}{1 - p} & \text{if } y_i = y_i \end{cases} \]

Then, it is optimal to take the action on states \( y \) such that:

\[ \sum_i \left[ \lambda_i + \tilde{\phi}_i(y_i) + (1 - \delta) \gamma_i(y) + \delta E[\gamma_i] \right] y_i > 0 \] (6)

**Proof.** See appendix. \( \square \)

This lemma does not provide a complete solution, as the Lagrange multipliers are still endogenous. However, it shows that the optimal decision making rule can be still regarded as a weighted voting rule. Indeed, we can rewrite the previous condition as a weighted voting rule, where the weights and the target depend on the current shock. Namely:

\[ m_i(y) = (\lambda_i + \delta E[\gamma_i])(y_i - \bar{y}_i) + \phi_i \bar{y}_i \frac{1}{1 - p} + (1 - \delta)(\gamma_i(y_{-i}, \bar{y}_i) - \gamma_i(y_{-i}, y_i, \bar{y}_i)) \] (7)

\[ M_i(y) = -\sum_i \left[ (\lambda_i + \delta E[\gamma_i])y_i - \phi_i \bar{y}_i \frac{p}{1 - p} + (1 - \delta)\gamma_i(y_{-i}, \bar{y}_i)\bar{y}_i \right] \] (8)

The Lagrange multipliers are state dependent, and therefore the voting weights are stochastic. On example 3, we will see that stochastic voting weights can be interpreted as a ‘probability of being in a council’ which resembles the UN voting system.

Note that the Pareto weights and the Lagrange multipliers affect similarly the voting weight. Thus, the self-enforcing requirement can be seen as a way to create a new profile of pseudo Pareto weights. Another remark is that heterogeneity of the members gives a different result from Maggi and Morelli
(2006). In their paper, the countries are homogeneous and therefore if the Pareto efficient cannot be implemented, the optimal mechanism is unanimity. However, when the countries are heterogeneous, it is possible to implement an outcome better than unanimity when \( \delta \) is smaller than, yet close to \( \delta^* \). This can be seen in the following result:

**Proposition 2.** Assume that \( \delta < \delta^* \), and recall the thresholds \( \delta_i \) defined in the proof of proposition 1. A necessary condition for the optimal equilibrium to implement a payoff higher than unanimity but lower than the Pareto optimum is that there is at least two countries \( i \) and \( j \) such that \( \delta_i < \delta < \delta_j \).

**Proof.** The first step is to show that there are at least two countries such that \( \delta_i < \delta_j \). Since \( \delta < \delta^* \), the first best cannot be implemented. Moreover, since the choice of \( x \) is discrete, from equation (4) the only way to provide incentives is by changing the second term on the left hand side, that is the expected payoff. This means that, by the definition of a Pareto allocation, any potential candidate for equilibrium will give to at least one country a payoff (strictly) smaller than the payoff from the Pareto efficient allocation. If \( \delta_1 = \delta_2 = \cdots = \delta_N \), then (by definition) these thresholds are all also equal to \( \delta^* \). Therefore, any outcome different from the Pareto efficient allocation will violate the participation constraint (4) for at least one country. As a consequence, there are at least two countries such that \( \delta_i < \delta_j \).

Next, we note that by the definition of \( \delta^* \), there is at least one country \( j \) such that \( \delta < \delta^* = \delta_j \). The last step is to show that there must be one other country such that \( \delta_i < \delta \). Let us assume that this does not hold. Therefore, \( \delta \leq \delta_i \) for all \( i \). This means that, in order to satisfy everyone’s participation, the candidate to optimal equilibrium must give every country a payoff higher than the Pareto efficient allocation. However, that is impossible.
The intuition for this is simple, if the discount factor is not large enough for implementing the Pareto efficient allocation, the organization has to provide a more favorable outcome to some countries. However, this means that some other countries will receive a payoff that is lower than the payoff they would get at the Pareto efficient allocation. In order for those countries to still be willing to participate from the decisions of the organization, a minimum requirement is that they have some slack in their participation constraint. That is, $\delta_i < \delta$. Moreover, if the threshold for all countries is the same ($\delta_1 = \delta_2 = \cdots = \delta_N$), it is not possible to ‘transfer’ some payoff from one country to another without violating the participation constraint. In particular, this is the case when countries are homogeneous.

An immediate corollary of this result is that there is a discontinuity in the organization’s value function:

**Corollary 1.** There is a discount factor $\tilde{\delta}$ satisfying $\min\{\delta_1, \delta_2, \ldots, \delta_N\} \leq \tilde{\delta} < \delta^*$ such that for any $\delta < \tilde{\delta}$, the best possible outcome is unanimity. Moreover, the organization’s value function is discontinuous at $\tilde{\delta}$.

*Proof.* See appendix.

Finally, it is important to recall that because the decision variable $x$ is restricted to be discrete the number of possible allocations is finite. As a consequence, there is a number of discontinuities on the value function between $\tilde{\delta}$ and $\delta^*$. The previous results are illustrated in figure 3.

### 2.7 Continuous Choice Variable

There is not much more to say regarding the optimal equilibrium when the choice variable is discrete. However, if we allow $x$ to take values on $[0, 1]$, we get some more interesting results. As we discussed on example 1, there are two ways to look at the continuous choice case. (i) The first option would be
randomization. That is, on a given state $y$, the action is implemented with probability $x(y)$. (ii) The second option would be to allow $x$ to be partially implemented. That is, countries compromise and only a fraction $x$ of the action is implemented. The second option adds some additional assumptions regarding the technology to implement the action as well as the preferences, however it essentially does not change the restrictions. Therefore we start by studying the partial implementation case:

**Partial Implementation**

Let us first study the simplest case. On the maximization problem (1), we allow $x$ to take any value on the interval $[0, 1]$. Moreover, since the the payoffs are stationary, nothing else changes, except that whenever equation 6 is an equality, it could be optimal to partially implement the action. Since we are expanding the set of choices, it is immediate that the expected payoff is higher. Moreover, the discontinuity shown in corollary 1 does no longer hold. The reason is that a key part of the proof in proposition 2 is that $x$ is discrete. Indeed, we can show that,

**Proposition 3.** If $x$ can take values on the interval $[0, 1]$: 
(i) There is always a payoff strictly better than unanimity for any $\delta > 0$,

(ii) The value function is continuous, and

(iii) The truth-telling condition in (3) is not binding at the optimum.

Proof. See appendix.

The intuition for this is simple. (i) Since the problem is linear, any $\delta$ admits one feasible allocation that dominates unanimity; namely a convex combination of the unanimity and the Pareto efficient allocations. (ii) Moreover, the number of possible allocations now is infinite, so the value function is continuous. (iii) Finally, by ignoring the truth-telling condition (3), the optimal rule is (weakly) increasing in $y$, therefore, from lemma 4, it still satisfies the truth-telling condition.\(^{10}\)

Randomizing Mechanism

On the previous section, $x \in (0,1)$ was regarded as if countries were able to compromise by partially implementing the action. Because of that assumption, the maximization problem essentially did not change. However, that imposed additional assumptions on the technology for implementing the action as well as the preferences. In this section, we still allow $x$ to take values in the interval $[0,1]$. Namely, we allow the possibility of randomization. However, the implementation of the action still requires to be either fully implemented or preserve the status quo. On each period, country members reveal their preferences, the organization randomizes and implements the action with some probability. If the outcome of the randomization dictates that the action is not taken, no incentive needs to be provided. However, if the outcome of the randomization dictates that the action is taken,

\(^{10}\)This was not necessarily true in the case when $x$ is discrete. However, even though $x$ was not is pointwise increasing in the discrete case, I could not find an example showing that the expectation of $x$ given $y_i$ was not increasing in $y_i$. 23
the participation constraint must be satisfied for \( x = 1 \). Because of this, the maximization problem has some significant changes. For the new maximization problem, \( x \) will denote the probability that the action is implemented and it can take any value on the interval \([0, 1]\). With this new interpretation, the relevant adjustment we need to do is on the participation constrain described in equation (4).

Thus we have the new following participation constraint:

\[
x_{\tau}(y_{\tau}) \left( (1 - \delta)y_{i,\tau} + \delta \left( (1 - \delta) \sum_{t=\tau+1}^{\infty} \delta^{t-\tau} \sum_{z_t \in Y} \Pr(z_t) x_t(z_t) z_{i,t} \right) \right) \geq 0, \forall i, y_{\tau} \tag{4'}
\]

In words, if with some positive probability the action is implemented, \( x_{\tau}(y_{\tau}) > 0 \), then the participation constraint needs to be satisfied. Moreover, note that so far the problem was linear, however because of this new participation constraint, the maximization problem is no longer linear. Note that any solution for the discrete choice case is feasible in this randomized mechanism. Moreover, any solution to the randomizing mechanism is feasible in the ‘partial implementation’ mechanism. Indeed,

**Proposition 4.** (i) The solution to the randomizing mechanism weakly dominates the solution to the discrete choice mechanism. However, both of them are dominated by the partial implementation mechanism.

(ii) The value function of the randomizing mechanism is discontinuous.

**Proof.** See appendix. \(\square\)

Note that when the Pareto efficient allocation is not feasible, in all the three mechanisms studied, there is randomness in the voting power, even in the discrete choice. Namely, in the discrete choice mechanism, the optimal voting weights were shock dependent. In addition, in the other two mechanisms, the decision making power depends on both: the current shock and a random variable that is
endogenous to the optimization of the problem.

2.8 Discussion and Extensions

2.8.1 Non Stationary Equilibrium

We saw that when the Pareto efficient allocation is the optimal stationary mechanism, it is also the optimal non-stationary (or unrestricted) mechanism. However, a stationary mechanism is not the best outcome when the countries are not very patient. In this section, I solve such unrestricted mechanism. To do this, I will follow two steps. First, I decompose the expected discounted payoff in two parts, a present payoff and a continuation payoff. The continuation payoff is the discounted sum of payoffs from next period onward, and it is history dependent. Second, I solve the maximization problem for each state of the world.

Step 1: Payoffs’ Decomposition

Let us rewrite the objective function in equation (1) as:

$$\sum_{y_0 \in Y} Pr(y_0) \sum_{i \in N} \lambda_{i,0} \left[ (1 - \delta)x_0(y_0)y_{i,0} + \delta \sum_{y_1 \in Y} Pr(y_1)w_{i,1}(y_1|y_0) \right]$$

Where $\lambda_{i,0} = \lambda_i$ and

$$w_{i,1}(y_1|y_0) = (1 - \delta)\sum_{t=1}^{\infty} \delta^{t-1} \sum_{y_t \in Y} Pr(y_t)x_t(y_t|y_1, y_0) \sum_{i \in N} \lambda_i y_{i,t}$$

is the expected discounted payoff of country $i$ in case that next period’s shock is $y_1$ and given that this period’s shock is $y_0$. 

25
Then, for a given state $y$ the simplified problem is:

$$\max_{x,w} \sum_{i \in N} \lambda_i \left[ (1 - \delta)xy_i(y) + \delta \sum_{z \in Y} Pr(z)w_i(z|y) \right]$$

s.t.

$$(1 - \delta)xy_i(y) + \delta \sum_{z \in Y} Pr(z)w_i(z|y) \geq 0$$

and

$$(w_i(z|y))_{i \in N} \text{ belongs to the set of self-enforceable equilibrium payoffs } W(z).$$

**Step 2: Maximization**

If I assign a Lagrange multiplier $\gamma_i$ to each action-taking incentive constraint, and denote $a_i(y)$ as the indicator function on the event $y_i = \bar{y}_i$, the problem becomes:

$$\max_{x,w_i,\gamma_i} \sum_{i \in N} (\lambda_i + \gamma_i) \left[ (1 - \delta)xy_i(y) + \delta \sum_{z \in Y} Pr(z)w_i(z|y) \right]$$

s.t.

$$\gamma_i \left( (1 - \delta)xy_i(y) + \delta \sum_{z \in Y} Pr(z)w_i(z|y) \right) = 0$$

$$(w_i(z|y))_{i \in N} \in W(z), \text{ and } \gamma_i > 0$$

Similarly to what I did before, the solution to this problem is equivalent to a weighted voting mechanism. The difference with the stationary case is that here the Lagrange multiplier plus the Pareto weight of period $t$ become the Pareto weight of period $t + 1$. Given that the mechanism
not only chooses today’s action but also continuation payoffs, it will choose equilibrium continuation payoffs weighted by \( \lambda_i + \gamma_i \). So, the mechanism is already implicitly using \( \lambda_{i,1} = \lambda_{i,0} + \gamma_i \) as the baseline Pareto weights for some implicit maximization problem that is going to happen in the future. Thus I have proved:

**Proposition 5.** There are two thresholds \( \delta^* > \delta^{**} \) such that for any \( \delta \geq \delta^* \) the optimal mechanism is stationary meaning that the voting weights are not state dependent and do not change over time. Moreover the payoffs are the same as in the Pareto efficient allocation. For \( \delta \leq \delta^{**} \) the unique solution is unanimity. For \( \delta^* > \delta > \delta^{**} \) the optimal mechanism is a weighted voting rule and the decision power changes over time. Countries who disagreed yet complied with actions that are beneficial to the organization will be rewarded in the future by a higher expected voting power.

There are two degrees in which countries can cooperate with each other. First, by coordinating in which states to take an action and in which ones to preserve the status quo, the members can attain efficient outcomes. However the voting weights that lead to this particular (efficient) outcome depend heavily on the Pareto weights, and choosing one of them may seem arbitrary. One way to endogenize these Pareto weights, and therefore the voting weights, is explained by the second degree in which countries can cooperate: the organization can increase countries’ (relative) voting weight in order to provide them incentives to participate in unfavorable states. So we can say that on top of choosing a decision rule that is correlated with the preferences, on a higher order of cooperation, the optimal mechanism rewards participation with future decision power.
2.8.2 Outside Options

Let us recall that I assumed that the payoffs after the organization breaks up is the same as the status quo forever. Instead, I can assume that each member has a different outside option \( b_i \). The introduction of an outside option does not greatly affect the equilibrium of the model, but it adds a relevant testable implication. The ‘Voluntary Membership’ constrain in equation 2 may be binding, and this would increase\(^{11}\) the voting weight of countries who may prefer to exert their outside option.

Let us take the stationary equilibrium to illustrate how outside options affect the voting power. Equation (5) is modified in the following way: the left hand size represents payoffs on the equilibrium path, so it does not change. However the right hand size depends on two terms. First, an instant payoff of not complying which is zero plus the discounted payoff of not having the organization, \( b_i \).

\[
(1 - \delta)y_i + \delta U_i \geq \delta b_i
\]  

(9)

Additionally, the voluntary membership restriction could be binding in this case. So, we need to take this restriction into account:

\[
U_i \geq b_i
\]

(10)

Let us use a Lagrange a multiplier \( \gamma_i(y) \) for the participation constrain of member \( i \) in state \( y \), and a Lagrange a multiplier \( \phi_i \) for the voluntary membership constrain of member \( i \). The objective function will be:

\(^{11}\)See equation (11) and the comment right below it.
\[
\max_{x, U_i} \sum_{y \in Y} Pr(y) \left[ \sum_{i=1}^{N} (\lambda_i + \gamma_i(y)) \left( (1 - \delta)x(y)g_i(y) + \delta(U_i - b_i) \right) \right] + \sum_i \phi_i(U_i - b_i)
\]

Since the sum of the probabilities adds up to one, we can introduce the term \( \sum_i \phi_i(U_i - b_i) \) inside the brackets. Moreover, the payoff \( U_i \) is nothing more than the expected payoff of member \( i \), so in equilibrium, it equals \( \sum_{y \in Y} Pr(y)x(y)g_i(y) \). After simplifying some terms, we obtain that the objective function is:

\[
\max_{x} \sum_{y \in Y} Pr(y)x(y) \sum_{i=1}^{N} \left[ (1 - \delta)(\lambda_i + \gamma_i(y)) + \delta \sum_{z \in Y} (\lambda_i + \gamma_i(z)) + \phi_i \right] y_i(y),
\]  

plus a constant independent of the decision variable \( x \). So, the voting weight:

\[
\left[(1 - \delta)(\lambda_i + \gamma_i(y)) + \delta \sum_{z \in Y} (\lambda_i + \gamma_i(z)) + \phi_i \right] (\bar{y}_i - y_i) \]

depends on the Lagrange multipliers of both restrictions. A larger outside option of country \( i \) increases the (equilibrium) lagrange multiplier \( \phi_i \) and therefore the voting weight of that country.

### 2.8.3 Imperfect Monitoring

So far, I focused on the case where the participation of each member is perfectly observable. In this section, I will briefly discuss what can happen when the organization receives imperfect signals of the members’ participation. This extension goes beyond the scope of the paper, so I will not provide a formal analysis.\(^{12}\) Moreover, the discussion here will be useful in the empirical test, Section 3.3.

Let us assume that the organization does not perfectly observe the participation of the members, but instead it receives imperfect signals of the compliance of each country. Namely, we can assume that there are two signals, one of the signals is more correlated with the country’s participation, and

\(^{12}\)For a formal derivation of a very specific example, please refer to Appendix A.
the other signal is more correlated with the country’s non-participation. Then, on every period, the
decision made by the organization \( x \) should depend also on the history of observed signals. Moreover,
in this case it is not necessarily optimal to use grim-trigger strategies (as the observed signals are
imperfect information regarding the participation of the members). In equilibrium, the expected
discounted payoff of a member should be higher after the organization has observed a signal that is more
correlated with the country’s participation in the collective action (good signal), and it should be lower
after the organization has observed a signal that is more correlated with the country’s participation in
the collective action (bad signal). This means that the voting weight of a country after a good signal
should be higher than after a bad signal.

3 Applications of the Model to the United Nations

Background

From all the International Organizations, one of the arguably most influential and powerful is the
United Nations. It is composed of several organs (such as the General Assembly and the Security
Council) and agencies (such as the IMF and the World Bank). According to their charter, the main
purpose of the United Nations is to maintain international peace and security. The organ devoted to
this specific task is the Security Council, which meets periodically to propose and vote on resolutions
that are compulsory\(^{13}\) to all members of the United Nations. However, only fifteen countries, five
permanent and ten non-permanent, have the right to vote at the Security Council (from a pool of
193 members). The ten non-permanent members have a tenure of two years, cannot be immediately

\(^{13}\)Country members are expected to follow the Security Council decisions, else they can receive sanctions from the UN itself.
reelected, and must win the elections by two thirds of the votes at the General Assembly.

In principle, this may suggest that there is some sort of inefficiency, as the preferences of a majority of the members is being ignored. Some questions that arise from this voting setup are: why and under which circumstances is it reasonable to ignore the opinion of the majority of the UN members? And, why would members comply with resolutions they did not even vote on? In this section I will apply the theoretical model to the voting system of the UN. First, I will use a numerical example to rationalize, within the framework of the model, the existence of the Security Council’s permanent members as well as their veto power. Second, I will briefly discuss three case examples that relate compliance with future veto power. Finally, I will test empirically two predictions of the model: compliance with unfavorable actions and outside option, both increase a country’s voting power and therefore its probability to be part of the Security Council.

3.1 Veto Power and Permanent Members

The historical explanation to the five permanent members is that the winners of WW2 decided to start an organization that seeks to prevent war and at the same time guarantee these countries’ power.\footnote{For example, see Bourantonis (2005).} However, this voting system could fit withing the framework of the model. We saw that a weighted voting rule can include veto power when the weight of a country is high enough. If the Pareto weights of the five permanent members (P5) were initially very high, that could explain the actual voting system of the UN. In the following numerical example, I compute the optimal stationary equilibrium for patient countries and for non-very patient countries. On the later case, I show that the optimal self-enforcing mechanism can assign veto power to a subset members. I also show that there can be different voting weights that implement the same outcome. Moreover, on particular way to do so is
by a council-like voting system, where only a subset of members can vote, but among such subset of members the votes are ‘equally’ valuable (one-country-one-vote).

**Example 3**

In this example, I will illustrate a case where: (i) the Pareto efficient outcome cannot be implemented, (ii) the mechanism gives veto power to some members, and (iii) other members vote only with some probability.

Let us assume that there are six countries: \{A, B, C, D, E, F\}. Moreover, only the first three members have a positive Pareto weight \((\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0, 0)\). There are two types of payoff supports. Member A has a payoff of either 3 or \(-8 + \epsilon\), and the rest of the members have a payoff of either 4 or \(-5\). Finally, the probability of being in favor of taking an action is \(p = 3/4\).

We will solve the equilibrium in a series of steps. On the first three steps, we will ignore incentives and focus on the efficient outcome:

**Step 1.** First, we use the notation of lemma 2 to compute the voting weights at the Pareto efficient allocation. Moreover, let us ignore the epsilon in the payoff of member A, and use it only in the case of being indifferent. The worst possible loss is: \(- \sum \lambda_i y_i = \frac{1}{3}8 + \frac{1}{3}5 + \frac{1}{3}5 = 6\). The voting weight of member \(i\) is \(\lambda_i(y_i - y_i)\), so we can compute the profile of voting weights: \((\frac{11}{36}, 3, 3, 0, 0, 0)\).

**Step 2.** Now we can determine (ignoring enforceability) when it is efficient to take the action. First we notice that only the first three members can vote. Also, from lemma 2, it is efficient to take the action when the sum of the weights of the members voting ‘yes’ exceeds the worst possible loss, which is six. This happens on the events in which at least two of the first three members
agree on taking the action. Notice that we use the epsilon being positive to break the tie in favor of taking the action on the events when $A$ disagrees, but $B$ and $C$ agree. Therefore, at the Pareto efficient allocation, only members $A$, $B$, and $C$ can vote. Also, looking only at the set of members who have the right to vote, the outcome is equivalent to one-country-one-vote with majority rule (or ‘egalitarian’ between countries $A, B, C$).

Step 3. Next, let us compute the Pareto efficient payoffs. For the first three members, their expected payoff is the sum of two terms: $p(p^2 + 2p(1 - p))\gamma_i + (1 - p)p^2\gamma_i$. In words, the first term is the probability of agreeing with the action, times the probability of implementing the action conditional on already having a ‘yes’ vote, times the payoff in such state. Similarly, the second term is the probability of disagreeing with the action, times the probability of implementing the action conditional on already having a ‘no’ vote, times the payoff in such state. The payoff of the last three members is simply the probability that the action is taken times the (unconditional) expected payoff, because they are never pivotal: $(p^3 + 3p^2(1 - p))(p\gamma_i + (1 - p)\gamma_i)$. The profile of payoffs (ignoring the term $\epsilon$) at the Pareto efficient allocation is: $(\frac{63}{64}, \frac{135}{128}, \frac{135}{128}, \frac{189}{128}, \frac{189}{128}, \frac{189}{128})$.

So far, I have computed the Pareto efficient outcome ignoring self-enforceability. So, the next steps will find the threshold $\delta^*$ that allow an efficient outcome, as well as show an example of an optimal stationary equilibrium with veto power, a voting council, and heterogeneous probabilities of voting.

Step 4. Given the profile of payoffs at the Pareto efficient allocation, we can check for what values of the discount factor $\delta$ the participation constraint is satisfied: $(1 - \delta)\gamma_i + \delta U_i \geq 0$. Each country has potentially a different threshold for the discount factor, and the first one to bind is the largest of them. The profile of such thresholds for the discount factor is: $(\frac{512}{575}, \frac{64}{91}, \frac{64}{91}, \frac{640}{829}, \frac{640}{829}, \frac{640}{829}) \simeq$
(0.89, 0.70, 0.70, 0.77, 0.77, 0.77). So, the actual threshold is the largest of them: \( \delta^* = \frac{512}{575} \approx 0.89 \).

Step 5. For the example, let us pick a discount factor smaller than the threshold \( \delta = 0.8 \). I solve the equilibrium numerically, and obtain the following decision rule:

We can classify the set of states \( Y \) in six categories:

1. Countries \( A, B, C \) agree.
2. Countries \( A \) and \( B \) agree, \( C \) disagrees, at least one of the countries \( D, E, F \) agrees.
3. Countries \( A \) and \( B \) agree, \( C \) disagrees, countries \( D, E, F \) all disagree.
4. Countries \( A \) and \( C \) agree, \( B \) disagrees, at least one of the countries \( D, E, F \) agrees.
5. Countries \( A \) and \( C \) agree, \( B \) disagrees, countries \( D, E, F \) all disagree.

On the optimal self-enforceable stationary equilibrium, it is optimal to take the action in states that fall into categories: 1, 2, and 4. And it is optimal to preserve the status quo in states that fall into categories 3, 5, and 6. This outcome can be implemented using a council-like voting system, for instance:

(i) There is a council of four members.

(ii) Countries \( A, B, \) and \( C \) are permanent members of the council.

(iii) Countries \( D, E, \) and \( F \) have an ex-ante probability of voting equal to \( 1/3 \) (but whenever the state of the world allows it, we must make sure that the one who is elected is going to vote in favor of taking the action).
(iv) Country A has veto power.

(v) The council passes resolutions if 3/4 of the members are in favor.

Alternatively, this can be implemented using voting weights: (0.75, 0.2, 0.2, 0.1, 0.1, 0.1) and a target of 1.\textsuperscript{15}

Step 6. To finalize the example, on figure 4, I show the optimal stationary mechanism payoffs for different values of $\delta$.

Figure 4: Sum of Expected Payoffs and $\delta$.

3.2 Examples of Compliance and Security Council Elections

Angola

The Angolan Civil War started in 1975 and lasted until 2002. The two major belligerents were the political parties MPLA (Movimento Popular de Libertação de Angola) and UNITA (União Nacional para a Independência Total de Angola). The leader of MPLA, José Eduardo dos Santos, was elected president in 1979, and held power since then. His counterpart was, Jonas Savimbi, the leader and

\textsuperscript{15}Moreover, there are other ways to implement the exact same outcome with other council sizes, as well as other voting weights. This is due to the fact that the number of Pareto efficient allocations (without allowing for randomization) is finite.
founder of UNITA. The UN Security Council has been increasingly bringing the Angolan Civil War into their agenda, and imposing sanctions towards UNITA. In early 2002, Savimbi died in combat against government troops; and without proper leadership, UNITA signed a cease of fire in that same year.

A quick inspection of the UN Security Council resolutions regarding Angola shows a contrast between the year 2002 and the years preceding it. For example, in 2001 there were three resolutions on that subject, all of them either reaffirming or requesting actions from both the Angolan government and UNITA, none of them showing satisfaction with the Angolan government’s compliance. In 2002, there were six resolutions regarding Angola. All of them had showed satisfaction with Angola’s compliance with UN requests. Every resolution has the word ‘welcoming’ on at least one of the paragraphs. This contrast shows satisfaction from the UN Security Council with Angola’s compliance managing their civil war in the year 2002. As a consequence, by the end of that same year, the elections for Security Council seats were held, Angola was elected, and became a non-permanent member in 2003-2004.

Rwanda

Rwanda is ethnically divided into Hutus, Tutsis, and Twas. In the early 90’s the government was in charge of Juvénal Habyarimana (a Hutu), from the MRND (Mouvement républicain national pour la démocratie et le développement) party. In 1990, the rival party RPF (Rwandan Patriotic Front) representing the Tutsis started a civil war that lasted until 1993. The UN Security Council did not have a strong position regarding Rwanda during this civil war, as there are no resolutions mentioning

---

17 Resolutions 1336, 1348, and 1374.
18 A typical resolution highlights the first word of each paragraph. This first word determines the tone of the paragraph, some examples are ‘welcoming’ to express satisfaction with the performance of a country; ‘requesting’ to express an action to be taken; ‘Reaffirming’ to remind of a decision previously made at an older resolution.
19 After signing the cease of fire, President Habyarimana was assassinated in 1994. This event lead to the Rwandan Genocide. See Fenton (2003).
Rwanda except in year 1993. In that year, there are four resolutions\textsuperscript{20} regarding Rwanda, and two of them show satisfaction with the Rwandan government’s compliance. By the end of that year, Rwanda was elected as a non-permanent member of the Security Council.

**Slovenia**

The case of Slovenia is different from the previous two examples. The argument I propose here will be linked to the rest of the former Yugoslav states. While in the two African cases previously explained the reward seemed to be ‘compliance for a Seat’, in this case the reward is not directly related to Slovenia’s compliance. Indeed, the election of Slovenia could have happened because all other former Yugoslav states had a poor compliance with UN resolutions\textsuperscript{21} and the UN needed a Yugoslav state at the Security Council as a representative for that region.

In 1991, a civil war started and Yugoslavia broke up into five states (Croatia, Bosnia and Herzegovina, FYR Macedonia, Serbia and Montenegro, and Slovenia) that later were admitted as members of the United Nations. However, the war continued until 2001. Moreover, none of these new countries gained the privilege of being part of the Security Council; and by 1997, the UN had not had a former Yugoslav representative since Yugoslavia’s last election in 1987.

Slovenia was elected as a non-permanent member of the Security Council by the end of 1997. In the years preceding its election, hardly any UN resolution mentions Slovenia. However, all other former Yugoslav states were often mentioned in UN resolutions regarding conflicts\textsuperscript{22}. For instance, in year 1997, there were several resolutions involving: Croatia, Bosnia and Herzegovina, FYR Macedonia, Serbia and Montenegro.

\textsuperscript{20}Resolutions 812, 846, 872, and 891.

\textsuperscript{21}Let us recall the discussion about imperfect monitoring. If the organization observes ‘bad signals’ of a country, it is punished by having a low voting power. If all the Slovenia’s neighbors have had a low compliance, then Slovenia’s voting power should increase in relative terms.

\textsuperscript{22}See Manusama (2007), chapters 2.3 and 4.3.
as well as other neighbor countries. All of these resolutions were either reaffirming decisions made in previous years (implying the relevance of old resolutions), or expressing concerns regarding the compliance of these three countries. Only one of those resolutions (Res. 1093) has the word ‘welcoming’ being addressed to the countries in conflict. However, that paragraph is addressed to all former Yugoslav states as a group. This suggests that lack of a Yugoslav state representative at the Security Council, plus the no compliance of Croatia, Bosnia and Herzegovina, and FYR Macedonia, lead to the election of Slovenia.

3.3 An Empirical Prediction

In this section, I will study the elections of the ten non-permanent seats at the Security Council. If we look at the historical data of the UN Security Council elections, we can see that some countries have been elected as non-permanent members more often than others.\(^{23}\) The observed outcome of the elections is generated from some distribution of voting power. It would be interesting to know what this distribution depends on.\(^{24}\) In order to do so, I test the two empirical predictions of the model. First, voting weights still depend on country characteristics, even when the voting weights are stochastic as in equation 7. Second, countries who complied yet disagreed with UN resolutions are expected to have a higher voting power in the future. Namely, they expect to have a higher probability of being part of the Security Council.\(^{25}\)

**What would we ideally want to test?** There are two types of variables that I am interested in. The first set of variables would capture country characteristics that influence the voting weights. In principle, this would include many factors such as the probability of having a conflict, the army

\(^{23}\)For example, the most often elected members and the number of years they held a seat are: Brazil (20), Japan (20), Argentina (18), Colombia (14), India (14), and Pakistan (14).

\(^{24}\)See the literature discussion related to Dreher and Vreeland (2014).

\(^{25}\)As shown in the numerical example, a council-like voting system can have the same outcome as voting weights.
size, the ability and speed to manufacture weapons, possible alliances, etc. The second set of variables would reflect the compliance of each country with each resolution, the position (preference) of each country regarding each of the resolutions, and how beneficial (or costly) it is for a country to implement such action.

The Econometric Model

I use a standard conditional logistic model.\textsuperscript{26} A group, \( k \), is defined as the subset of countries competing for Security Council seats at a particular moment; that is, a group is a pair (region, year) and there are five regions Africa, Asia-Pacific, Latin America and Caribbean, Eastern Europe, and Western Europe and others. For instance, this means that the elections for African seats in the year 2002 did not have anything to do with the elections in the Asia-Pacific region in the same year. Even though this does not perfectly capture the way the theoretical model works\textsuperscript{27} it is probably close to what we see in the actual UN elections.

A group has a number of observations \( L_k \). The number of observations is the number of eligible countries. An eligible country in group \( k \) is any member who belongs to the region in \( k \), and has not been in the Security Council for the past two years. This excludes current members (because tenure is for two years) and members who just left the Security Council (because this is a UN rule).

On each group \( k \), the outcome of (an eligible) country \( l \) equals: \( r_{k,l} = 0 \) if country \( l \) has not been elected, and \( r_{k,l} = 1 \) if the country has been elected. Let \( T_k = \sum_l r_{k,l} \) be the number of members elected in group \( k = (\text{region, year}) \). The probability of observing a profile of outcomes \( r_k = (r_{k,1}, r_{k,1}, \ldots, r_{k,L_k}) \)

\textsuperscript{26}Chamberlain (1980).

\textsuperscript{27}As seen in Example 3, even if the voting weights are independent, the probabilities of being part of the council may be correlated.
given the number of elected members $T_k$ is:

$$Pr(r_k|T_k) = \frac{\exp \left( \sum_{l=1}^{L_k} r_{k,l} V_{k,l} \beta \right)}{\sum_{s_k \in S_k} \exp \left( \sum_{l=1}^{L_k} s_{k,l} V_{k,l} \beta \right)}$$

where $S_k$ is the set of all possible combinations of outcome profiles that have $T_k$ elected members, $V_{k,l}$ are the explanatory variables, and $\beta$ is vector of parameters. Finally, the conditional log-likelihood is:

$$\log L = \sum_k \left( \sum_{l=1}^{L_k} r_{k,l} V_{k,l} \beta - \log \left( \sum_{s_k \in S_k} \exp \left( \sum_{l=1}^{L_k} s_{k,l} V_{k,l} \beta \right) \right) \right)$$

**Description of the Available Data**

**Data on compliance:** Unfortunately, there is hardly any data on the compliance of countries with Security Council resolutions. The only data on compliance that is publicly available is from the International Peace Institute (IPI), but they only limit attention to resolutions on civil war from 1989 to 2003.\(^{28}\) Moreover, out of the 193 UN members, only 37 countries are involved in such conflicts (27 had a civil war in that period, and 10 were directly mentioned in the resolutions). Additionally, a civil war is not the ideal type of shock to test the model, as it involves only a few neighbor countries in addition to the one at civil war.

IPI’s data lists each resolution, the action requested, and the country addressed.\(^{29}\) For each one of these cases, the variables that I will use are: the date of the resolution, a dummy indicating whether the country agreed or disagreed with the resolution, the difficulty level of the action demanded by the UN (which is ranked from one to three), and finally the level of compliance which is ranked in

---

\(^{28}\) The database extends to 2006, but the data coded after 2003 does not include any measure of compliance.

\(^{29}\) In general, each resolution has a number of requests to several countries.
four categories: none or low, medium-low, medium-high, high to full. Moreover, the compliance is measured at two different moments. There is a short-term compliance and a mid-term compliance. The short-term compliance is measured within six months after the resolution was passed. The mid-term compliance is measured two years after the resolution was passed.

**Data on country characteristics:** I use four variables from the World Bank database: real gdp, population, land size, and military expenditure.

**Data on Security Council Elections:** I use UN data of Security Council elections. A country that starts working as a non-permanent member of the Security Council in a given year \( t + 1 \) was elected in the last trimester of the year \( t \). Therefore, Elections on the year \( t \) are based on variables of the same year: country characteristics and compliance. Finally, a country is eligible if it is neither a current member of the Security Council nor has been part of it the previous year.

**Construction of the Compliance Rate**

In this section, I describe how I construct a compliance rate using the data from IPI. Let us recall that the compliance level of each resolution is measured twice: short-term (6 months after adoption) and mid-term (2 years after adoption) compliance. Each of them is classified in four categories: *none or low*, *medium-low*, *medium-high*, *high to full*. I want to assign values to each one of these categories in such a way that a positive score is given to ‘good behavior’ and a negative score is assigned to ‘bad behavior.’ Moreover, since I do not know how the UN actually values each of the four levels of compliance coded by IPI, I proceed as follows: *(i)* I assign a low value to *none or low compliance*, *(ii)* each improvement in compliance receives a constant incremental value, *(iii)* the two lower compliance levels have a value below zero and the two higher ones are above zero, *(iv)* and the support of the
values is symmetric around zero. That is:

\[ \text{comp}_{\text{medium,low}} - \text{comp}_{\text{none,or,low}} = \text{comp}_{\text{medium,high}} - \text{comp}_{\text{medium,low}} = \text{comp}_{\text{high,full}} - \text{comp}_{\text{medium,high}} \]

and

\[ \text{comp}_{\text{medium,low}} + \text{comp}_{\text{medium,high}} = 0 \]

Following these rules, I assign values of \(-3, -1, 1, 3\) respectively, where the constant increment is 2.\(^{30}\)

The next step is to compute a compliance score for each country, action, and resolution. I weight each level of compliance by the difficulty level\(^{31}\) which goes from 1 to 3. Additionally, I also weight each level of compliance by the prior consent of each country; being zero weight\(^{32}\) if the country agreed with the resolution, and a weight of one if the country disagreed with the resolution.\(^{33}\) That is because, if a country agrees with the action, the participation constraint should not be binding. So, for each resolution, request, and country addressed, the compliance score (either short or mid term) is:

\[ \text{comp.score} = \text{comp}_{\text{level}} \times \text{difficulty} \times \text{prior.consent} \]

From the way this compliance score is constructed, countries who agreed with a resolution will get a score of zero on such war. Similarly, a country who is not involved in a given civil war will also get

---

\(^{30}\)This constant increment is just a normalization. Additionally, I could also make these values ‘convex,’ by making the increases in these values larger when going moving from the middle to the extremes, for instance \(-4, -1, 1, 4\). This would translate in higher punishments and higher rewards as the levels are further away from the middle. Similarly, the values could be ‘concave,’ for instance \(-2, -1, 1, 2\). However, when running the regressions, the two alternatives mentioned here do not greatly affect the results.

\(^{31}\)This difficulty level is already coded as 1,2 or 3 by IPI.

\(^{32}\)Following the model, if a country agreed, the participation constraint is not binding.

\(^{33}\)IPI coded two additional categories: ‘no agreement but an allied agreed’, and ‘no agreement but an enemy agreed.’ I consider them as ‘yes’ and ‘no’ respectively. However, I also tested the result computing these two categories both as ‘yes’ and both as ‘no’, and it did not change the results in a significant way, as the proportion of these two categories is small in the data set.
a score of zero on such war. Finally, I add up each of these scores by country by year, and define this sum as the compliance rate of country $i$ in period $t$: Short-term compliance rate and Mid-term compliance rate.$^{34}$

---

$^{34}$For an example of what the compliance rate looks like, see the appendix, table 4.
Table 3: UN Security Council Elections as a function of Compliance and Country Characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Baseline Model; No Military Spending; Only Compliance; Only Country Characteristics; Ignoring Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security Council Election</td>
<td></td>
</tr>
<tr>
<td>short-term compliance</td>
<td>1.263*</td>
</tr>
<tr>
<td></td>
<td>(0.147)</td>
</tr>
<tr>
<td>mid-term compliance</td>
<td>0.809*</td>
</tr>
<tr>
<td></td>
<td>(0.0835)</td>
</tr>
<tr>
<td>log(gdp)</td>
<td>1.759*</td>
</tr>
<tr>
<td></td>
<td>(0.419)</td>
</tr>
<tr>
<td>log(pop)</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>(0.176)</td>
</tr>
<tr>
<td>log(land)</td>
<td>0.903</td>
</tr>
<tr>
<td></td>
<td>(0.0944)</td>
</tr>
<tr>
<td>log(military)</td>
<td>1.038</td>
</tr>
<tr>
<td></td>
<td>(0.205)</td>
</tr>
<tr>
<td>Observations</td>
<td>1339</td>
</tr>
<tr>
<td>Group</td>
<td>Region, Year</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The outcome variable is the probability of election at the Security Council. This table reports the odds-ratio; standard errors in parentheses. The odds of a country to be elected is the ratio of the probability of election and the probability of non-election. The odds-ratio represents the effect that a marginal increase in an explanatory variable has on the odds of a country to be elected. If the odds-ratio is larger than one, then the variable has a positive effect on the odds of election.
Results of the Empirical Test

The sample size is from 1989 to 2003, and I include all UN members. The explanatory variables that I am interested in are the short-term compliance rate and country characteristics. I include the mid-term compliance rate as a placebo test.\textsuperscript{35} Table 3 shows the results for different specifications of the model. To interpret the results, let us notice that the coefficients displayed are the odds ratio. An explanatory variable with an odds-ratio higher than one has a positive effect on the outcome of elections. For example, the short-term compliance in the baseline model has an odds-ratio of 1.263. This means that if a country increases its compliance rate by one unit, the odds (probability of election divided by probability of no-election) of that country increases by about 26%.

Compliance in the short-term has a positive effect on the elections for seats at the Security Council. This goes along with what the model would predict: a country who disagreed with a resolution but still participated with the decision made by the organization will be rewarded with higher voting power in the future. On the other hand, the effect that the mid-term compliance rate has on the probability of election is not positive, which a good indicator of not having a spurious regression.\textsuperscript{36} Finally, from the country characteristic, only GDP is significant. The result that only GDP is significant among country characteristics is similar to Dreher and Vreeland (2014).\textsuperscript{37} Therefore, compliance and country characteristics have the expected effect on the Security Council elections.

\textsuperscript{35}The mid-term compliance from IPI is coded according to compliance two years after the resolution was passed. This means that it was not observable at the moment of elections. Therefore, while the short-term compliance is expected to have a positive effect on the Security Council elections, the mid-term compliance is expected not to have such positive effect.

\textsuperscript{36}The mid-term compliance rate has a negative and significant coefficient. A possible explanation for this is that countries could comply in the short-term at high costs in order to show good behavior until the UN elections happen, and then shirk by deteriorating the compliance. Indeed, another version of the model was tested separating the mid-term compliance in two variables: mid-term-increases with respect to short-term and mid-term-decreases with respect to short term. The former is not significant, and the later has a significant and negative sign. However, further analysis is beyond the scope of this paper.

\textsuperscript{37}They include country characteristics as well as a "turn-taking" variable, which measures how long a country has not been part of the security council. They find that GDP and Population are significant in some regions, and the turn-taking variable is significant for all regions.
4 Conclusions

I have studied the optimal decision rule for an International Organization. This optimal decision rule induces an endogenous weighted voting system. When the members are patient, the Pareto efficient allocation can be implemented and therefore the voting weights are constant. When the members are not very patient, the decision power is a random variable. If the organization is under a stationary equilibrium, the distribution of voting power is constant over time. However, if the organization is following a non-stationary equilibrium, the distribution of voting power is history dependent. In the later case, countries who participate by taking actions in unfavorable states are rewarded by future decision power. Finally, I have argued that, under imperfect monitoring, signals correlated with compliance have a positive effect on the voting weight, and analogously signals correlated with no-compliance have a negative effect on the voting weight.

The UN has currently a voting system which has some properties that resemble the optimal voting rule when the members are not very patient. I have shown with a relatively simple example a case where the best stationary equilibrium can be implemented using a council-like voting system with: heterogeneous probabilities of voting among members and veto power for a subset of the council. Finally I tested two predictions of the model. I found that the only country characteristic that is significant for the Security Council elections is GDP. This would suggest that a country’s income is the only relevant country characteristic. This can be due to the fact that a rich country can create alliances and afford military. I also found that compliance with the Security Council resolutions that is observable at the moment of elections increases the probability of a country to win a seat as a non-permanent member. Both results go along with the theoretical model.
Appendix

A  An example of Imperfect Monitoring

In this section, I will study a relatively simple information structure; this is just an example to show how imperfect signals can affect the voting weights. I will assume that there is imperfect monitoring only on one of the members, say member $N$. Let $d \in \{0, 1\}$ be this member’s decision to participate. The organization observes a signal $\tilde{d} \in \{0, 1\}$ with the following distribution: the probability of observing a signal $\tilde{d} = d$ given that the actual decision was $d$ is $q > 1/2$. So, if this member participates, there is a chance $1 - q < 1/2$ that the organization will observe a signal $\tilde{d} = 0$ which is more associated with no-participation. Finally, I will assume that the outcome is not observable, however the outcome is still deterministic in the sense that if everyone participates, the action is implemented. This means that the members only observe the participation of $N - 1$ members and the imperfect signal of member $N$; and based on that, they get a sense of their payoff, but do not actually observe it. Moreover, if the equilibrium is such that the members are expected to comply with the decisions of the organization, we can say that the members know their payoff, even if they do not observe it.

For simplicity, let us focus on the stationary equilibrium. Let $\tilde{U}_i(\tilde{d})$ be the expected payoff of member $i$ if the signal of last period was $\tilde{d}$. If we define $U_i = q\tilde{U}_i(1) + (1 - q)\tilde{U}_i(0)$, the incentive constraint of members 1 to $N - 1$ does not change from equation (5). However, member $N$’s decision has an effect on the distribution of the signal $\tilde{d}$. The participation constrain for such member is:

$$(1 - \delta)xy_N + \delta \left( q\tilde{U}_N(1) + (1 - q)\tilde{U}_N(0) \right) \geq \delta \left( q\tilde{U}_N(0) + (1 - q)\tilde{U}_N(1) \right)$$

This is the equation that guarantees that member $N$ participates by taking the action whenever
it is asked to do so. The left hand side is the equilibrium payoff: the payoff in the current period is 
\((1 - \delta)xy_N\), and the continuation payoff depends on the signal. However, what is different from before is the right hand side. Since the organization does not observe whether country \(N\) complied or not, the payoff of not taking the action is not necessarily zero. If country \(N\) does not take the action, it knows that the current period’s payoff is zero, but with probability \(q\) the realized signal is \(\tilde{d} = 0\) and therefore the expected future payoff is \(\tilde{U}_i(0)\), and with probability \(1 - q\) the realized signal is \(\tilde{d} = 1\) and therefore the expected future payoff is \(\tilde{U}_i(1)\). I will assign a Lagrange multiplier \(\gamma(y, \tilde{d})\) to the incentive constrain when the current state is \(y\) and the signal observed on last period is \(\tilde{d}\). If we proceed as in section 2.8.2 (assign Lagrange multipliers, and simplify the objective function), we obtain that the voting weight for member \(N\) is the following:

\[
\left( (1 - \delta)q(\lambda_N + \gamma_N(y, 1)) + \delta \left( \lambda_N + E \left[ \gamma_N(y, \tilde{d}) \right] \right) (2q - 1)(1 + \delta(2q - 1)) \right) (y_N - y_{\bar{N}})
\]

and

\[
\left[ (1 - \delta)(1 - q)(\lambda_N + \gamma_N(y, 0)) - \delta \left( \lambda_N + E \left[ \gamma_N(y, \tilde{d}) \right] \right) (2q - 1)(1 + \delta(2q - 1)) \right) (y_N - y_{\bar{N}})
\]

for a period following a good signal \((\tilde{d} = 1)\) and a bad signal \((\tilde{d} = 0)\), respectively. Notice that the first expression is larger than the second one. This means that there is a drop in the voting weight of member \(N\) on the period after observing a bad signal. When there is imperfect monitoring, the organization provides incentives by rewarding the members who drew a signal that is more correlated with compliance, and by punishing the members who drew a signal that is more correlated with no-compliance. If we think of the implications of imperfect monitoring in the UN, we can say that
the probability of a member to be elected to the Security Council increases with higher evidence of compliance.

B  Example of Compliance Rate

A country’s compliance score of a resolution is the product of the compliance level times the difficulty of the task times the country’s consent on that resolution. The compliance rate of a country in a given year is the sum of all the scores of such country in the selected year. The following table shows the non-zeros compliance rates in the year 1998.

Table 4: Compliance rate of countries involved in a civil war, 1998

<table>
<thead>
<tr>
<th>Short-Term Compliance Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afghanistan               -35</td>
</tr>
<tr>
<td>Angola                    -9</td>
</tr>
<tr>
<td>Bosnia and Herzegovina    4</td>
</tr>
<tr>
<td>Georgia                   -4</td>
</tr>
<tr>
<td>Guinea                    -1</td>
</tr>
<tr>
<td>Haiti                     -7</td>
</tr>
<tr>
<td>Namibia                   -4</td>
</tr>
<tr>
<td>Rwanda                    1</td>
</tr>
<tr>
<td>Senegal                   -5</td>
</tr>
<tr>
<td>Sierra Leone              3</td>
</tr>
<tr>
<td>Tajikistan                1</td>
</tr>
</tbody>
</table>
References


