

# A Minimax Procedure for Negotiating Multilateral Treaties

Steven J. Brams  
Department of Politics  
New York University  
New York, NY 10003  
USA  
[steven.brams@nyu.edu](mailto:steven.brams@nyu.edu)

D. Marc Kilgour  
Department of Mathematics  
Wilfrid Laurier University  
Waterloo, Ontario N2L 3C5  
CANADA  
[mkilgour@wlu.ca](mailto:mkilgour@wlu.ca)

M. Remzi Sanver  
Department of Economics  
Istanbul Bilgi University  
80310, Kustepe, Istanbul  
TURKEY  
[sanver@bilgi.edu.tr](mailto:sanver@bilgi.edu.tr)

Prepared for delivery at the 2004 Annual Meeting of the American Political  
Science Association, September 2 – September 5, 2004. Copyright by the  
American Political Science Association.

### **Abstract**

A procedure for reaching agreement on multilateral treaties, based on “fallback bargaining,” is proposed. The compromise it finds minimizes the maximum distance, called the Hamming distance, between it and the top preferences of all players. This compromise may differ from the compromise produced by the usual procedure—voting on each treaty provision—which minimizes the sum of distances. The proposed procedure is relatively invulnerable to strategizing, inducing players to be truthful in expressing their preferences.

The application of the procedure requires that issues be of more or less equal significance to countries and that they be as independent as possible. Applying the procedure to oil-pollution control negotiations among 32 countries in 1954 yields six compromise outcomes, all different from that produced by the usual procedure. Approval voting is suggested as a way to break ties among the compromise outcomes.

*Keywords:* Multilateral treaty; dispute resolution; fallback bargaining; voting; environment.

## A Minimax Procedure for Negotiating Multilateral Treaties

### 1. Introduction

In this paper we propose a procedure for reaching agreement on multilateral treaties that produces a compromise as close as possible to the preferences of all parties. By “close” we mean that the maximum distance of the compromise from the position of any state is minimal, which we call a *minimax outcome*. We show that this procedure is relatively invulnerable to strategizing by states, reducing any incentive they might have to misrepresent their preferences to try to induce a better outcome.

The procedure is different from the usual method for reaching an agreement in multilateral treaty negotiations. Normally states vote separately on each provision of a treaty, often starting from a “single negotiating text.”<sup>1</sup> A provision is included in the treaty if a simple or qualified majority—or sometimes all states—supports it; otherwise, it is excluded. If provisions are included in a treaty that are not approved by all states, then those states that do not get their way on provisions they consider important might be quite frustrated—and unwilling either to sign the treaty or to abide by its terms.

Thereby a large majority of states may agree to and achieve a treaty far removed from the ideal of one or more states. By contrast, if the votes for and against every provision are aggregated in a different way, a better compromise—one that leaves no state too aggrieved—may be found.

Our focus is in on aggregating votes, which is appropriate when negotiations occur in the shadow of voting. That is, the parties to a treaty, when they decide what provisions to propose, must consider how they will be viewed when it comes time to vote on them.

---

<sup>1</sup> The extensive use of such a text in the Law of the Sea negotiations is discussed in Sebenius (1984).

Will they receive sufficient support to be accepted by most if not all parties in the negotiation? Even if a provision is found acceptable, will its inclusion facilitate or retard passage of the treaty in a referendum, or in a parliamentary vote, in those countries that must ratify it?

Clearly, voting—or, more accurately, its anticipation—is part and parcel of negotiating a treaty, especially in influencing what gets put in and what gets left out as well as how the treaty provisions are packaged. Although we do not model the crafting and packaging of these provisions, the fact that the minimax procedure we propose does not hurt any party as much, potentially, as the usual voting procedure suggests that it would allow the parties to be more open in their negotiations. To be sure, we have no evidence to prove this, but we do present evidence that a 1954 environmental treaty might have fared better if the positions of countries had been aggregated under the minimax procedure rather than the usual procedure.

To begin the analysis, we suppose that negotiations have reached a stage whereby the provisions of a treaty in dispute (i) can be specified, (ii) are of approximately equal significance to all states, and (iii) are relatively independent of each other. While states of very different size, wealth, military capability, and the like can be weighted (perhaps by monetary contributions, as in the International Monetary Fund, or IMF), the procedure we propose does not directly reflect such differences.

An approach for addressing these differences is to group states into regional or functional blocs that are more or less equal in size, possess substantial common interests, and have a similar stake in the outcome of the treaty negotiations. Putting together such blocs, however, is not straightforward, as we discuss later.

The procedure we propose for forging consensus in multilateral treaty negotiations is based on “fallback bargaining” (Brams and Kilgour, 2001), but we modify it in an important way. Instead of assuming that states directly rank alternatives, we represent their preferences by vectors of 1’s and 0’s, where a “1” indicates approval and a “0” indicates disapproval of each proposed provision of the treaty.

We assume that all positions of states, given by vectors of 1’s and 0’s, are possible. Each state supports (“votes for”) a specific vector, which we call that state’s *top preference*, or ideal point. Below a state’s top preference, its preference for other vectors depends on their distance from its ideal: It next most prefers any of the vectors that differ from its top preference in one component, then in two, and so on, down to the vector that differs from the top preference in every component. Thereby each state ranks vectors, based on their proximity to its ideal point.

For example, assume that only two provisions of a treaty are at issue, and a state approves of the first provision and disapproves of the second. Then its top preference is 10; its ranking in our model is

$$10 > \{11, 00\} > 01.$$

That is, the player most prefers 10, next most prefers either 11 or 00—between which it is indifferent—and least prefers 01.

The plan of the paper is as follows. In section 2 we describe and illustrate fallback bargaining, showing that an outcome may be highly dependent on the decision rule used. Indeed, in one example we show that every possible outcome may be selected under some fallback *decision rule*  $r$ , which may range from 1 to all  $n$  players.

We focus on *fallback bargaining with unanimity* ( $FB_n$ ), comparing it with the common procedure of majority voting (MV) on each provision. We show that the minimax outcomes under  $FB_n$  may differ from MV outcomes, which we call *minisum outcomes* because they minimize the sum (or average) distance to players' top preferences. Whereas  $FB_n$  outcomes are in the Rawlsian tradition of minimizing the largest deviations from a settlement, MV outcomes are in the utilitarian tradition of minimizing average departures.

In section 3 we compare minimax and minisum outcomes, asking which is better and under what conditions. We introduce the notion of a “weighted” minimax outcome to take into account the fact that the weight or the number of players having the same top preference may vary if negotiation is not among equals.

In section 4 we analyze the manipulability of MV and  $FB_n$ , showing that  $FB_n$  is vulnerable to manipulation; this is also true if the fallback decision rule is not unanimity. By contrast, MV is invulnerable—players always have an incentive to be truthful. But in any realistic situation with incomplete information about the preferences of two or more players,  $FB_n$  would be extremely difficult to manipulate.

$FB_n$  outcomes seem superior in situations like the one discussed in section 5, wherein 32 states negotiated oil-pollution controls in 1954. In these negotiations, states could abstain as well as vote yes or no on treaty provisions. To break ties among the six  $FB_n$  outcomes, we suggest approval voting, which allows states to approve of one or more of the  $FB_n$  outcomes and which is quite resistant to strategic exploitation.

We conclude in section 5 by suggesting that  $FB_n$  may well facilitate consensus in multilateral treaty negotiations, especially those that include most of the nearly 200 states

in the world today. These complex negotiations frequently involve both individual states and overlapping blocs of states, scores of provisions, and considerable maneuvering by the players to try to achieve a strategic advantage. We believe that our proposed procedure would encourage players to be honest, render their negotiations more open, and make the compromises they achieve as acceptable as possible to all players.

## 2. Fallback Bargaining under Different Decision Rules

Assume there are  $k$  provisions of a treaty being negotiated by  $n$  players (countries). A possible treaty is a binary  $k$ -vector,  $(p_1, p_2, \dots, p_k)$ , where  $p_i$  equals 0 or 1. Such binary vectors will be called *combinations*. To simplify notation, we write combinations such as  $(1, 1, 0)$  as 110. Note that the total number of combinations is  $2^k$ .

A country's *top preference* is its most-preferred combination. We assume countries rank treaties according to their "Hamming distance" from their top preferences. The *Hamming distance*,  $d(p, q)$ , between two binary  $k$ -vectors,  $p$  and  $q$ , is the number of components on which they differ. For example, if  $k = 3$  and a player's top preference is 110, the distances between it and the eight binary 3-vectors (including itself) are shown below:

<b>Top Preference</b>	<b><math>d = 0</math></b>	<b><math>d = 1</math></b>	<b><math>d = 2</math></b>	<b><math>d = 3</math></b>
110	110	100 010 111	000 101 011	001

Note that three vectors are tied for second place at distance  $d = 1$ , and three more are tied for third place at  $d = 2$ . In general, the preference ordering is not strict because of such ties.

To find a consensus choice, we next describe fallback bargaining for some decision rule  $r \leq n$ :<sup>2</sup>

1. Assume countries approve only of combinations at distance  $d = 0$  from their top preferences—that is, only the top-preference combinations are acceptable. If one or more combinations is approved of by at least  $r$  countries, then those with the most approvals are winning, and the process stops.

2. If no combination is approved of by at least  $r$  countries at distance  $d = 0$ , then consider all combinations at distance  $d \leq 1$  from the players' top preferences. If one or more combinations is approved of by at least  $r$  countries, then those with the most approvals are winning, and the process stops.

3. If no combination is approved of by at least  $r$  countries at distance  $d \leq 1$ , continue the descent until, *for the first time*, one or more combinations is approved of by at least  $r$  countries at distance  $d \geq 2$ . The combinations with the most approvals are winning, and the process stops.

Exactly when the process stops depends on, among other things, the number of countries  $n$ , the number of provisions  $k$ , and the decision rule  $r$  (Brams and Kilgour, 2001). Our first result characterizes winning outcomes under  $\text{FB}_n$ .

---

<sup>2</sup> Technically, Brams and Kilgour (2001) define fallback bargaining only when preferences form a linear order, which assumes that players have strict preferences. But it is easy to extend this procedure to nonstrict preferences, as we do here.

**Proposition 1.** *The  $FB_n$  winners are the minimax outcomes—they minimize the maximum distance to the top preference of any player.*

**Proof.** See Brams and Kilgour (2001, p. 292, Theorem 3).

The idea behind the proof is the following. Suppose an  $FB_n$  outcome is not a minimax combination. Then there is some other combination for which the maximum distance to the top preference of any player is less. But the descent under  $FB_n$ , which stops at the *first* point at which all players approve of some combination, must stop at this other combination. Therefore, this other combination must be a minimax combination.

A decision rule of unanimity is frequently used to decide important questions, such as the admission of new members into a regional or international organization. As a case in point, the Treaty of Rome in 1958 made unanimous consent of the original 6-member Common Market a requirement for admission of new members; that rule is still in effect in the present 15-member European Union that has evolved from the Common Market over the last 45 years.

That different fallback decision rules can give dramatically different results is illustrated by Example A, in which there are  $n = 10$  players and  $k = 3$  provisions, which

---

—

Example A about here

---

—

yield 8 combinations. Note that four combinations (000, 100, 010, 001) are the top preferences of one player each, and three combinations (110, 101, 011) are the top

preferences of two players each; only combination 111 is nobody's top preference.

Geometrically, the positions of the different players are shown in 3-space in Figure 1.

---

Figure 1 about here

---

We use Example A to prove our next proposition.

**Proposition 2.** *It is possible for every combination to win, or tie for winning, under some decision rule  $r$ .*

**Proof.** The number of players that approve of each combination at distances  $d = 0, 1, 2,$  and  $3$  are shown in Example A. At  $d = 0$ , this number is simply the number of players whose top preference is that combination. The three combinations that are approved of by 2 players (110, 101, 011) are the winners under decision rules  $r = 1$  and  $r = 2$ : For  $r = 1$ , they get the most votes (2); for  $r = 2$ , they are the only combinations that get 2 votes.

It is easy to verify that four combinations (100, 010, 001, 111) are the winners at distance  $d = 1$  under decision rules  $r = 3, 4, 5,$  and  $6$ ; and one combination (000) is the winner at distance  $d = 2$  under decision rules  $r = 7, 8, 9,$  and  $10$ .<sup>3</sup> In sum, every combination is a winner, or tied for winning, under some decision rule  $r$ . Q.E.D.

Because the winning combination under fallback with unanimity ( $FB_n$ ), 000, is the only combination that is a distance of 2 or less from the top preference of every player, it

would seem a reasonable compromise. In Figure 1, this position is a distance of  $d = 3$  from combination 111, but because 111 has no players supporting it, combination 000 is a maximum distance of  $d = 2$  from all players.

Majority voting (MV) on each provision is probably the procedure most often used to select a winning combination. As we next show, the outcome it produces may not coincide with the  $FB_n$  winner.

**Proposition 3.**  *$FB_n$  may produce a unique outcome that is different from an MV outcome and is no player's top preference.*

**Proof.** In Example B, there are  $n = 5$  players and  $k = 4$  provisions, which yield 16

---

—

Example B about here

---

—

combinations, whose numbers we show in parentheses following the combination. Combination 0010 (4) is the unique winner under  $FB_n$  at  $d = 2$ . But this is not the top preference of any of the 5 players, whose top preferences are combinations 1000 (2), 0001 (5), 0110 (9), 0011 (11), and 0111 (15) (shown in boldface in Example B).

It can be verified directly that a majority of the players prefer 0 on the first two components, and 1 on the last two components, so combination 0011 (11) is the MV winner. Hence, a unique  $FB_n$  winner may not be the top preference of any player and also may differ from an MV winner. Q.E.D.

---

<sup>3</sup> Clearly, all players must approve of all combinations at  $d = k$ , which is  $d = 3$  in Example A. But because all players find combination 000 acceptable at  $d = 2$  in this example, it is the unanimity winner ( $r = 10$ ).

Because the  $FB_n$  winner, combination 0010 (4), is selected at  $d = 2$ , its maximum distance from all players' top preferences is 2. By contrast, the MV winner, combination 0011 (11), is  $d = 3$  from combination 1000 (2)—the top preference of one of the players—so it is not as close to *all* players as the  $FB_n$  winner, which is what distinguishes  $FB_n$  winners from MV winners, as we will discuss in more detail shortly.

In Example A,  $FB_n$  gives (uniquely) combination 000, whereas MV gives all 8 combinations (because positions of 0 and 1 tie with 4 votes each on all 3 provisions). In this case, some tie-breaking mechanism would have to be used to select a specific winning combination.

When  $FB_n$  and MV give different outcomes, does one yield a “better” compromise than the other? Consider again Example B, wherein we noted that  $FB_n$  gives combination 0010 (4) and MV gives combination 0011 (11). What explains this difference?

Observe that the distances of the five players from  $FB_n$  winner 0010 (4) are (2, 2, 1, 1, 2); the maximum distance is 2 and the sum of the distances is 8. By contrast, the distances from MV winner 0011 (11) are (3, 1, 2, 0, 1); the maximum distance is 3 and the sum of the distances is 7. As shown in Proposition 1,  $FB_n$  minimizes the maximum distance; as shown next, MV minimizes the sum of distances.

**Proposition 4.** *The MV winners are the minisum outcomes—they minimize the sum of the distances to the top preferences of all players.*

**Proof.** Assume that there are  $k$  provisions and  $n$  players, and that player  $i$  supports the binary  $k$ -vector  $p^i = (p_1^i, p_2^i, \dots, p_k^i)$ . For an arbitrary binary  $k$ -vector  $x = (x_1, x_2, \dots, x_k)$ , define

$$d_j(x, p^i) = \begin{cases} 0 & \text{if } x_j = p_j^i \\ 1 & \text{if } x_j \neq p_j^i \end{cases}$$

for  $i = 1, 2, \dots, n$  and  $j = 1, 2, \dots, k$ . Then it is clear that the Hamming distance from  $x$  to

$p^i$  is given by  $d(x, p^i) = \sum_{j=1}^k d_j(x, p^i)$ . We consider how to select  $x$  so as to minimize

$$D(x) = \sum_{i=1}^n d(x, p^i).$$

For any  $x$  and  $j$ , define  $S_j(x) = \sum_{i=1}^n d_j(x, p^i)$ . Note that  $D(x) = \sum_{j=1}^k S_j(x)$ , and that

$S_j(x)$  depends only on  $x_j$  and not on the other  $k - 1$  components of  $x$ . From the definition of  $d_j(x, p^i)$  it follows that

$$S_j(x) = \begin{cases} N_j(0) & \text{if } x_j = 1 \\ N_j(1) & \text{if } x_j = 0 \end{cases}$$

where  $N_j(t) = |\{i: p_j^i = t\}|$  for  $t = 0$  and  $1$ . It also follows that  $x$  minimizes  $S_j(x)$  iff (if and only if)  $x_j = 1$  when  $N_j(0) < N_j(1)$  and  $x_j = 0$  when  $N_j(0) > N_j(1)$ . [Note that when  $N_j(0) = N_j(1)$ , there is no condition on  $x_j$ .] Consequently,  $x$  minimizes  $D(x)$  iff

$$x_j = \begin{cases} 1 & \text{if } N_j(1) > N_j(0) \\ 0 & \text{if } N_j(0) > N_j(1) \end{cases}$$

for all  $j = 1, 2, \dots, k$ , which is true iff  $x$  is an MV outcome. Q.E.D

The idea underlying the proof of Proposition 4 is that the MV outcome minimizes the number of disagreements of the players on each provision, because at least as many players agree as disagree with the MV outcome on each provision. Because the sum of disagreements across the  $k$  provisions equals their sum across all the players, the latter—

which is the total distance of all players from their top preferences (Hamming distance)—is also minimized.

### 3. Which Minimization Criterion Is Best?

Is it better to minimize (i) the sum of the distances from a compromise outcome (MV winner) or (ii) the maximum distance of players from a compromise outcome ( $FB_n$  winner), given there is a conflict? If the goal is to avoid antagonizing any player “too much,” there are good grounds for choosing  $FB_n$ . In Example B, the  $FB_n$  winner is a maximum distance of 2 from the top preferences of the 5 players, whereas the MV winner is a distance of 3 from one player’s top preference.

Consider how these winning outcomes would (or would not) change if the players were differently weighted. Suppose, for example, that the player supporting combination 1000 (2) were given a weight of 5, or was replaced by 5 players that all have combination 1000 (2) as their top preference. Now this combination would become the MV outcome—on each provision, it would get majority approval—but the top preference of the player supporting combination 0111 (15) is a distance  $d$  of 4 from it [because combination 0111 (15) is the *antipodal combination*—the opposite on every component—of combination 1000 (2)].<sup>4</sup>

By contrast, the  $FB_n$  winner would still be combination 0010 (4), which remains at a distance  $d$  of at most 2 from each player’s top preference. However, now the sum of distances of each player from combination 1000 (2) is

$$5(2) + 2 + 1 + 1 + 2 = 16,$$

---

<sup>4</sup> Necessary and sufficient conditions under which minimax and minisum outcomes may be antipodal are given in Brams, Kilgour, and Sanver (2004).

whereas the sum of distances from MV combination 1000 (2) is

$$5(0) + 2 + 3 + 3 + 4 = 11,$$

which is considerably less, so combination 1000 (2) becomes the MV winner.

Patently, the players supporting combination 1000 (2) “call the shots” under MV. By implementing their top preference, they leave the minority—the other 4 players—at distances  $d$  of 2 to 4 from their top preferences.

Because  $FB_n$  depends only on *which*  $k$ -vectors are the top preferences of one or more players—not the numbers that support each—it is insensitive to weights (sometimes referred to as independence of “clones,” or other players with the same top preference). This property suggests that  $FB_n$  would work best for *negotiation among equals*, which might require grouping countries into blocs.<sup>5</sup>  $FB_n$  might also be used within a bloc to find a compromise position that best reflects the views of the bloc’s members.

If negotiation is not among equals, then it is possible to amend  $FB_n$  to take into account the different numbers or different weights of players supporting each combination. As an illustration, consider Example A, in which three combinations (5, 6, and 7) are supported by two players each, four combinations (1, 2, 3, and 4) by one player each, and one combination (8) by no player. Let the speed of descent from a top

---

<sup>5</sup> This is decidedly not the case in international forums like the UN General Assembly. But should a microstate have the same weight as the US or China in international negotiations? Requiring states to coalesce in blocs that have more or less equal weight and take a common position is one way to ensure negotiation among equals. In fact, the 184 countries in the IMF are grouped into 24 so-called Constituencies—with from 1 member (eight biggest contributors) to 25 members (all African states)—each having a weight equal to the total weight of its members. Each Constituency is represented by an Executive Director, who casts the Constituency’s votes as a bloc on an Executive Board that, in principle, reflects a consensus among the Constituency’s members. The decision rules used are 50+%, 70%, and 85%, depending on the importance of the issue being decided.

preference be inversely proportional to the number of players who support it, as long as this number is not zero.

Thus in Example A, the two players each that support combinations 5, 6, and 7 will consider acceptable combinations at  $d = 1$  exactly when the one player each that supports combinations 1, 2, 3, and 4 consider acceptable combinations at  $d = 2$ . In effect, each of the former combinations is the top preference of twice as many players as the latter combinations, producing twice as much “inertia” in the descent. Accordingly, the winning combination may be only half as distant from the top preferences of these players as from the top preferences of the players whose top preference is supported only by themselves. It is easy to see that the first combination to become acceptable to all 10 players under these rules is 111 (combination 8), which is  $d = 1$  from the 6 players that approve of two provisions each and  $d = 2$  from the 4 players that approve of one provision each.

Note that combination 111 is the antipode of the minimax outcome, 000. Combination 111 might be considered a *weighted minimax outcome* in the sense that it minimizes the maximum weighted distance from all players, where weight in this case reflects the numbers of players who make each combination their top preference. Because there is not a consensus in most treaty negotiations on how players should be weighted, we will not try to apply the notion of weighting here.

Many international disputes are between two countries, or two sets of countries. For example, the EU, as a collectivity, now negotiates with the US. In such two-player disputes, wherein the obvious compromise is to settle about half of the disagreements in favor of each of the two players,  $FB_n$  is hardly needed. On the other hand, if the

disagreements concern issues that the two sides care differently about, a point-allocation procedure called “adjusted winner” seems a better mechanism to use (Brams and Taylor, 1996, 1999).

#### 4. Manipulability

A bargaining procedure is *manipulable* if a player, by misrepresenting its preferences, can obtain a preferred outcome. To determine what is “preferred,” we must take account of preferences between sets. We assume that if  $a$  is preferred to  $b_1$  and  $a$  is preferred to  $b_2$ , then  $a$  is preferred to  $\{a, b_1, b_2\}$ .

Because both MV and  $FB_n$  require that players support a specific binary  $k$ -vector representing their top preferences, the manipulability of these procedures depends on whether the players can benefit by indicating a top preference different from their sincere one.

**Proposition 5.**  *$FB_n$  is manipulable, whereas MV is not.*

**Proof.** We start with MV. Because the players’ choices are binary on each provision, it is always in a player’s interest to support whichever position (0 or 1) it prefers. Since the majority choice on each provision has no effect on the majority choices on other provisions, each player has a dominant strategy of voting sincerely on all provisions.

Now consider  $FB_n$ . In Example C, there are  $n = 4$  players and  $k = 4$  provisions,

---

—

Example C about here

---

–

which yield 16 combinations, but not all are shown. For the “true top preferences” of the 4 players in Example Ca, there are three  $FB_n$  winners [combinations 1000 (1), 0000 (5), and 1001 (6)], each of which gets unanimous approval at distance  $d = 2$ .

If the player supporting combination 1000 (1) in Example Ca misrepresents its top preference as combination 1110 (5) in Example Cb, then the *unique* winner is combination 1000 (1), which is in fact this player’s true top preference (and which also happens to be the MV winner). Thus, by indicating an insincere top preference, this player can induce a preferred outcome in Table Cb, based on our assumption about preferences between sets (given at the beginning of this section). Q.E.D.

In theory, therefore,  $FB_n$  is vulnerable to manipulation. In practice, however,  $FB_n$  is probably almost as resilient to manipulation as MV is, because its possible exploitation would require that a manipulative player have virtually complete information about the preferences of other players—and that they all act sincerely—which is unlikely in most real-life situations.

The resilience of  $FB_n$  as well as MV to manipulation would seem to negate the advantages that accrue to strategic players in negotiation with naïve players. In fact, even for strategic players, there may be no gain from manipulation.

To illustrate, assume that the fallback decision rule in Example C is simple majority (3 of the 4 players) rather than unanimity, which we indicate by  $FB_m$ . As shown in Example Ca, the  $FB_m$  outcome is combination 1000 (1). It is not difficult to verify that

no player can induce a preferred outcome under  $FB_m$  by misrepresenting its top preference.

But manipulation is certainly possible under  $FB_m$ .

**Proposition 6.** *If the decision rule is not unanimity, fallback bargaining is manipulable.*

**Proof.** If the decision rule is not unanimity, we prove this proposition by giving an example that shows that fallback bargaining is manipulable under  $FB_m$ . In Example Da,

---

—

Example D about here

---

—

observe that 5 players choose 4 different combinations. Combination 0000 (1) receives the approval of 3 of the 5 players at distance  $d = 1$  and is the unique winner under  $FB_m$ .

Note that the most dissatisfied players will be the 2 players whose top preference is the antipodal combination, 1111 (4). If these 2 players misrepresent their top preference as 1100 (5) in Example Db, they induce combinations 1000 (2), 0100 (3), and 1100 (5) as the winners, each of which receives the approval of 4 players at distance  $d = 1$ .<sup>6</sup> Because these three winning combinations are closer to the true top preference of the 2

---

<sup>6</sup> Because combination 0000 (1) receives approval from 3 players, four combinations listed in Example Db receive approval from a majority of players. However, we assume that the  $FB_m$  winners are those combinations—namely, the three combinations, 1000 (2), 0100 (3), and 1100 (5)—that receive the most approvals.

disassembling players [combination 1111 (4) in Da] than is combination 0000 (1), their misrepresentation is rational, rendering  $FB_m$  manipulable.<sup>7</sup> Q.E.D.

We next turn to a real-life dispute that suggests how  $FB_n$  might be used in practice. It raises new issues (the possibility of abstention; ties among  $FB_n$  outcomes) for which we suggest some pragmatic solutions.

### 5. An Application to an Environmental Dispute

On April 25, 1954, a three-week conference of 32 states, representing 95 percent of the world's shipping tonnage, convened in London. It included 18 developed states from Europe, North America, and Australasia; of the others, 3 were from Eastern Europe, 4 from Asia, 6 from Latin America, and 1 from Africa.

The treaty that resulted from this conference, called the International Convention for the Prevention of Oil Pollution (OILPOL '54), prohibited the discharge of oil from certain ships in specified ocean areas. It went into force in 1958 and was followed by three more conferences between 1958 and 1962 that resulted in the adoption of additional measures to strengthen those in OILPOL '54.

The principal question on the table concerned how best to prevent oil pollution of the sea from discharges by both tankers and nontankers.<sup>8</sup> The debate focused on

a number of contentious issues: the seriousness of the problem; past experience

---

<sup>7</sup>  $FB_m$  is called the "majoritarian compromise" in Hurwicz and Sertel (1999), Sertel and Sanver (1999), and Sertel and Yilmaz (1999), wherein it is analyzed as a voting procedure with majority rule rather than a bargaining procedure, as in Brams and Kilgour (2001). But the fallback process is essentially the same under both interpretations, whether the decision rule is simple majority, qualified majority, or unanimity.

<sup>8</sup> The account that follows, and the voting of the 28 states shown in Table 5 (no data were given for 4 states), is taken from M'Gonigle and Zacher (1979, pp. 85-91). A detailed discussion of compliance with this and subsequent oil-pollution treaties can be found in Mitchell (1994). For data on the preferences of states for another treaty, see Hug and König (2002).

with [prohibition] zones, the technical feasibility of suggested preventive measures, and cost . . . [There was] a parallel between the alternative a state favored at the meeting and the information they seemed to possess on the persistence, behavior, and effects of oil in the marine environment . . . . In the main . . . it was differing perceptions of cost/benefit that dictated the choice between a general prohibition and a system of zonal controls (M'Gonigle and Zacher, 1979, pp. 86-87).

The 12 different positions of both individual states and subsets of states on four issues are shown in Example E. The first issue indicates a state's position on the

---

—

Example E about here

---

—

desirability of a general prohibition on discharges, based on its statements for or against such a prohibition. The other three issues are votes on specific resolutions: (i) forbidding discharges by tankers; (ii) extending the prohibition zone off the British Isles; and (iii) making the North Sea a prohibition zone.

Besides the positions of 1 (yes) and 0 (no) on these issues, we have included "A" to indicate either no statement or a vote of abstention by a state—in effect, no announced position, or a position of neutrality.<sup>9</sup> In measuring distance, we assume A to be a distance of zero from both 0 and 1. Thus, A indicates agreement, or at least acquiescence, with both the pro and con positions.

As can be seen from Example E, the positions of the states range from highly pro-environment (1111) to highly anti-environment (0000). Because a significant number of states chose A on one or more issues, neither 1 nor 0 wins by a majority on any of the four issues—nor, for that matter, does A.

On the first two issues, the plurality winner is 0, with this position getting the support of 14 states on each issue. On the second two issues, the plurality winner is A (13 players support this position on the third issue, and 14 states on the fourth issue). Consequently, 00AA (10) is the MV outcome if it can include A's.

Because A signifies either the lack of a position or neutrality, it is questionable whether it should be counted as a component of an outcome. Ruling out A as a component, we ask whether 0 or 1 would obtain the most votes on the third and fourth issues. It turns out that 0 would win on the third issue (with the support of 8 states) and 1 would win on the fourth issue (with the support of 9 states). Thus, if A is not permitted as a component of an outcome, the MV outcome would be 0001.

The latter outcome is different from the six binary outcomes under  $FB_n$ :

1100;    1001;    1001;    0110;    0101;    0011.

Each of these outcomes is a maximum distance of  $d = 2$  from the 12 positions of the 28 states shown in Table E, including the two extreme positions of 1111 (1) and 0000 (12).<sup>10</sup>

---

<sup>9</sup> Five states did not vote on resolution (iii); we presume their positions to be A even though they did not formally abstain.

<sup>10</sup> It is worth noting that position AAAA (8) is distance  $d = 0$  from *all* positions, including combinations not shown in Example E, and therefore would be the unique  $FB_n$  outcome if A were permitted. We think it proper to exclude this “compromise” as well as all other combinations that have any A's as components. Because they are failures to act, either by approving or not approving a provision that addresses some issue, they are not compromises at all. In our view, a multilateral treaty on the environment—or anything else—should be a statement of what will and will not be done on salient issues, not a deferral of action, or cop-out.

Observe that these compromises all involve two 1's and two 0's, putting them half-way between the extreme positions. By Proposition 1, they minimize the maximum distance separating them from the positions of all states, which is  $d = 2$ . By Proposition 4, the MV outcome, 0001, minimizes the sum of the distances (or the average distance) to all states. Notice that the MV outcome is a greater distance ( $d = 3$ ) from position 1111 (1) than is any of the six  $FB_n$  outcomes.

The  $FB_n$  outcomes seem more defensible as compromises than MV outcome 0001, in part because of the greater distance of the latter outcome from pro-environment position 1111 (1) that is supported by 5 states.<sup>11</sup> While the size and importance of the 28 states shown in Example E states is very different, when we compare the 14 states that lie above neutral position AAAA (8) with the 13 states that lie below this position, both subsets comprise similar mixes of large and small states, suggesting that the six  $FB_n$  outcomes offer a tolerable balance that reflects the different sizes, as well as the different interests, of all states.

If the six  $FB_n$  outcomes are considered more compelling as compromise choices than the single MV outcome, the question of which one of the former should be chosen remains open. To make a choice, we recommend *approval voting* (Brams and Fishburn, 1983), whereby states can approve of as many of the  $FB_n$  outcomes as they like. The outcome that receives the most approval is the winner, with ties broken randomly.

In the Example E, suppose that the states are demanding and approve only of  $FB_n$  outcomes that are a distance of  $d = 0$  from their positions. Then combination 0011 would

---

<sup>11</sup> Whether the  $FB_n$  outcomes would have resulted in greater compliance with OILPOL '54, which was spotty at best (Mitchell, 1994), is difficult to say. Subsequent treaties, especially those that raised equipment standards, did lead to greater compliance, but it is not evident that a more pro-environmental treaty in 1954 would have had this effect.

win, garnering approval from the 6 states that take positions AAAA (8), A0AA (9), and 00AA (10). If the states are more forgiving and approve of  $FB_n$  outcomes that are a distance of  $d \leq 1$  from their positions, then combination 0011 would win again, but so would combinations 1001 and 0101, all tying with approval from 15 states.

If the states are still more forgiving and approve of  $FB_n$  outcomes that are a distance of  $d \leq 2$ , the six  $FB_n$  outcomes would all tie, receiving the approval of all 28 states. Because the latter strategy admits all outcomes as acceptable, rational players would presumably be more discriminating in order that they influence the choice of an outcome.

While approval voting seems a good way for players to narrow down the set of compromise outcomes if several tie under  $FB_n$ , like all voting procedures it is vulnerable to manipulation.<sup>12</sup> But how likely are ties? There were none in Example B, Example Cb, and Example Da, in each of which the number of players was a significant proportion of the number of combinations (at least 25%).

It is not unreasonable to suppose that treaty negotiations may occur over, say, 20 provisions, in which case there would be approximately 1.05 million combinations. If 200 states all had different top preferences, there would be less than 1 state for every 5,000 combinations, making the probability that an  $FB_n$  winner is unique exceedingly high for most assumptions about the distribution of top preferences.<sup>13</sup>

---

<sup>12</sup> Approval voting is invulnerable if the preferences of voters are dichotomous, in which case voters have a dominant strategy of voting sincerely (Brams and Fishburn, 1983). More generally, Condorcet winners are strong Nash-equilibrium outcomes under approval voting, but other outcomes may be (nonstrong) equilibrium outcomes as well (Brams and Sanver, 2003).

<sup>13</sup> In the case of voting by 1.8 million Los Angeles County voters on 28 propositions (268.4 million combinations) in 1990, the ratio of voters to combinations was much less (1 voter for every 149 combinations). Still, the most votes received by any single yes-no combination was only 0.20% (1 out of every 500 voters), and that was for the combination recommended by the *Los Angeles Times* (Brams, Kilgour, and Zwicker, 1997, 1998). Without such coordination, the probability of ties seems negligible if

A more difficult question in applying  $FB_n$  is how to write provisions of a treaty so that they are (i) of approximately equal significance to all states and (ii) relatively independent of each other. If these criteria are not satisfied, then the assumption that a state's ranking of combinations is inversely related to the distance of these combinations from its top preference is not tenable. Drafting treaty provisions that satisfy these criteria is a delicate art that will surely require considerable intellectual effort and substantial good will on the part of the states, or blocs of states, negotiating a treaty.

## 6. Conclusions

We proposed a procedure, based on fallback bargaining, that we believe could help states reach consensus in negotiating and agreeing on multilateral treaties. While this procedure directly addresses voting rather than the negotiation of treaty provisions, we suggested that because the crafting and packaging of treaty provisions is done in anticipation of the voting procedure that will be used—and outcomes likely to occur under it—negotiation is difficult to separate from voting.

Because the procedure we proposed selects an outcome that is not too distant from the preferences of any party, it is likely to encourage a more frank exchange of views in the negotiation phase. While the procedure is manipulable in theory, in practice it is likely to be difficult to do so because of the complexity of calculations required to make manipulation successful under it.

---

$FB_n$  is applied to a significant number of treaty provisions. Supporting evidence in a different context can be found in the 2003 election by the Game Theory Society (GTS) of 12 new council members from a list of 24 candidates. Of the 161 GTS members who voted, which is about the maximum number of countries that participate in treaty negotiations, only two voters cast an identical ballot. It is worth noting that even though 16.8 million possible combinations (ballots) had to be checked, it was not difficult to determine the minimax outcome with the aid of the computer (Brams, Kilgour, and Sanver, 2004), indicating that the application of  $FB_n$  to 25 or so treaty provisions is feasible.

This procedure and the usual procedure for voting on treaty provisions begin with a set of proposed provisions. Under fallback bargaining with unanimity ( $FB_n$ ), they should be as equal in importance, and as independent of each other, as possible.

Each state indicates a top preference—by either approving or not approving each of the provisions—but abstentions are possible, as we showed. From the combination that is a player's top preference, a preference order is induced, based on proximity as measured by the Hamming distance: Combinations that differ by one component are ranked next-highest, combinations that differ by two components next, and so on down to the combination that differs on every component from a player's top preference.

$FB_n$  selects combinations that minimize the maximum distance of players from the compromise outcome(s) it selects (minimax outcome). If there are several such outcomes, as there were in the empirical example we analyzed, approval voting seems a good procedure to break ties among them. As we illustrated, the compromise selected may be quite different from the outcome that majority voting (MV) on each provision, which minimizes the sum of distances of players from a compromise outcome (minisum outcome), would select.

Insofar as the players, or blocs of players, are more or less equal, an  $FB_n$  compromise seems to us appealing. If the players or blocs are unequal in size or importance, this will not be reflected in the standard  $FB_n$  outcome; for this reason, we suggested a way to modify the procedure so as to reflect these different weights via different rates of descent. By contrast, MV takes weight into account, but it may yield outcomes that are compromises in name only—they may be far from the top preferences of some players.

While unanimity is sensible in many bargaining situations, a decision rule of simple or qualified majority decision rule may be more practicable in certain situations. Instead of requiring that the distance of all players from a compromise outcome be minimized, it would allow some players to be more distant as long as a (qualified) majority is as close as possible to the compromise outcome.<sup>14</sup> As with  $FB_n$ ,  $FB_r$  ( $r = m$  or a qualified majority) is vulnerable to manipulation, but fallback-bargaining procedures are likely to be difficult to manipulate in most practical situations.

While no bargaining procedure satisfies all desirable properties,  $FB_n$  seems to us a promising way to promote consensus in complex negotiations involving several states or blocs of states. In the environmental treaty negotiations we analyzed, it yielded a compromise that seemed superior to both the MV and  $FB_m$  outcomes. Besides its technical properties,  $FB_n$  is practicable and transparent, which should help to instill confidence that the outcome it gives respects, insofar as possible, everybody's interests.

---

<sup>14</sup> In the Example E,  $FB_m$  gives combinations 0100 and 0000 (12) with 16 approvals at  $d = 1$ , which are distances of 3 and 4, respectively, from combination 1111 (1), the top preference of 5 states.

### Example A

#### Number of Players Approving of Each Combination at Different Distances

<b>Combination</b>	<b><math>d = 0</math></b>	<b><math>d = 1</math></b>	<b><math>d = 2</math></b>	<b><math>d = 3</math></b>
1. 000	1	4	10***	10
2. 100	1	6**	8	10
3. 010	1	6**	8	10
4. 001	1	6**	8	10
5. 110	2*	4	9	10
6. 101	2*	4	9	10
7. 011	2*	4	9	10
8. 111	0	6*	9	10
<b>Total</b>	10	40	70	80

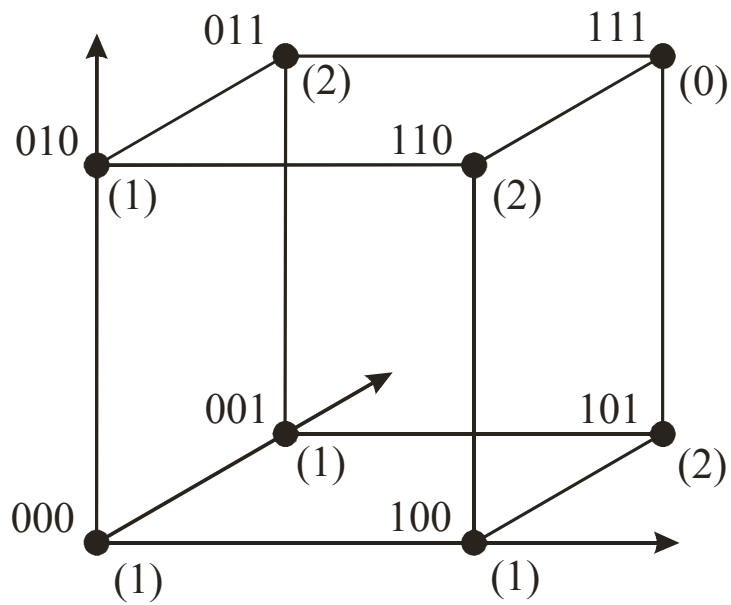
\*Winner for  $r = 1$  and 2.

\*\*Winner for  $r = 3, 4, 5,$  and 6 (simple majority).

\*\*\*Winner for  $r = 7, 8, 9,$  and 10 (unanimity).

Figure 1

Positions of  $n = 10$  Players on  $k = 3$  Provisions (Example A)



*Note:* The numbers in parentheses give the numbers of players supporting the corresponding top preference.

### Example B

#### Number of Players Approving of Each Combination at Different Distances

(Top Preferences of Five Players in Boldface)

Comb.	$d = 0$	$d = 1$	$d = 2$	$d = 3$	$d = 4$
1. 0000	0	2	4	5	5
<b>2. 1000</b>	<b>1</b>	<b>1</b>	<b>2</b>	<b>4</b>	<b>5</b>
3. 0100	0	1	4	5	5
4. 0010	0	2	5*	5	5
<b>5. 0001</b>	<b>1</b>	<b>2</b>	<b>4</b>	<b>5</b>	<b>5</b>
6. 1100	0	1	2	4	5
7. 1010	0	1	3	5	5
8. 1001	0	2	3	4	5
<b>9. 0110</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>5</b>	<b>5</b>
10. 0101	0	2	4	5	5
<b>11. 0011</b>	<b>1</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>5</b>
12. 1110	0	1	3	4	5
13. 1101	0	1	4	5	5
14. 1011	0	0	3	5	5
<b>15. 0111</b>	<b>1</b>	<b>3</b>	<b>4</b>	<b>4</b>	<b>5</b>
16. 1111	0	1	3	5	5
<b>Total</b>	<b>5</b>	<b>25</b>	<b>55</b>	<b>75</b>	<b>80</b>

\* $FB_n$  winner; the MV winner is combination 0011 (11).

### Example C

#### Number of Players Approving of Each Combination at Different Distances

##### Ca. True Top Preferences

Combination	$d = 0$	$d = 1$	$d = 2$
1. 1000	1	3**	4*
2. 0001	1	1	2
3. 1100	1	2	3
4. 1010	1	2	3
5. 0000	0	2	4*
6. 1001	0	2	4*
10 other comb.'s	0	At most 2	At most 3

##### Cb. Player 1 Above Misrepresents as Player 5 Below

Combination	$d = 0$	$d = 1$	$d = 2$
1. 1000	0	2	4*
2. 0001	1	1	1
3. 1100	1	2	3
4. 1010	1	2	3
5. 1110	1	3**	3
11 other comb.'s	0	At most 1	At most 3

\* $FB_n$  winner.

\*\* $FB_m$  winner.

### Example D

#### Number of Players Approving of Each Combination at Different Distances

##### Da. True Top Preferences

Combination	$d = 0$	$d = 1$
1. 0000	1	3*
2. 1000	1	2
3. 0100	1	2
4. 1111	2	2
12 other combinations	0	At most 2

##### Db. 2 Player 4's Above Misrepresent as 2 Player 5's Below

Combination	$d = 0$	$d = 1$
1. 0000	1	3
2. 1000	1	4*
3. 0100	1	4*
4. 1111	0	0
5. 1100	2	4*
11 other combinations	0	At most 2

\* $FB_m$  winner.

### Example E

#### Positions of 28 States on Four Different Issues at the 1954 Convention on the Prevention of Oil Pollution of the Sea

Positions	No. of States	States
1. 1111	5	Australia, Germany, Ireland, New Zealand, UK
2. 1A11	1	The Netherlands
3. 111A	1	India
4. 1AA1	2	Canada, Israel
5. 11AA	2	Poland, USSR
6. 10A1	1	Sweden
7. 01AA	2	Brazil, Portugal
8. AAAA	1	Venezuela
9. A0AA	1	Mexico
10. 00AA	4	Chile, Greece, Italy, Yugoslavia
11. 000A	3	France, Spain, US
12. 0000	5	Belgium, Denmark, Finland, Japan, Norway

MV outcome (A not permitted): 0001.

FB<sub>n</sub> outcome (A not permitted): 1100; 1001; 1010; 0110; 0101; 0011.

## References

- Brams, Steven J., and Peter C. Fishburn (1983). *Approval Voting*. Cambridge, MA: Birkhäuser Boston.
- Brams, Steven J., and D. Marc Kilgour (2001). "Fallback Bargaining." *Group Decision and Negotiation* 10, no. 4 (July): 287-316.
- Brams, Steven J., D. Marc Kilgour, and M. Remzi Sanver (2004). "A Minimax Procedure for Electing Committees." Preprint, Department of Politics, New York University.
- Brams, Steven J., D. Marc Kilgour, and William S. Zwicker (1997). "Voting on Referenda: The Separability Problem and Possible Solutions." *Electoral Studies* 16, no. 3: (September): 359-377.
- Brams, Steven J., D. Marc Kilgour, and William S. Zwicker (1998). "The Paradox of Multiple Elections." *Social Choice and Welfare* 15, no. 2 (February): 211-236.
- Brams, Steven J., and M. Remzi Sanver (2003). "Voter Sovereignty and Election Outcomes." Preprint, Department of Politics, New York University.
- Brams, Steven J., and Alan D. Taylor (1996). *Fair Division: From Cake-Cutting to Dispute Resolution*. Cambridge, UK: Cambridge University Press.
- Brams, Steven J., and Alan D. Taylor (1999). *The Win-Win Solution: Guaranteeing Fair Shares to Everybody*. New York: W. W. Norton.
- Hug, Simon, and Thomas König (2002). "In View of Ratification: Governmental Preferences and Domestic Constraints at the Amsterdam Intergovernmental Conference." *International Organization* 56, no. 2 (Spring): 447-476.
- Hurwicz, Leonid, and Murat R. Sertel (1999). "Designing Mechanisms, in Particular for

Electoral Systems: The Majoritarian Compromise.” In Murat R. Sertel (ed.), *Economic Design and Behaviour*. London: Macmillan.

M’Gonigle, R. Michael, and Mark W. Zacher (1979). *Pollution, Politics, and International Law: Tankers at Sea*. Berkeley, CA: University of California Press.

Mitchell, Ronald B. (1994). *Intentional Oil Pollution at Sea: Environmental Policy and Treaty Compliance*. Cambridge, MA: MIT Press.

Sebenius, James K. (1984). *Negotiating the Law of the Sea*. Cambridge, MA: Harvard University Press.

Sertel, Murat, and M. Remzi Sanver (1999). “Designing Public Choice Mechanisms,” in Imed Limam (ed.). *Institutional Reform and Development in the MENA Region*. Cairo, Egypt: Arab Planning Institute, pp. 129-148.

Sertel, Murat, and Bilge Yilmaz (1999). “The Majoritarian Compromise Is Majoritarian-Optimal and Subgame-Perfect Implementable.” *Social Choice and Welfare* 16, no. 4 (August): 615-627.