ECONOMIC RESEARCH REPORTS

Fallback Bargaining

by
Steven J. Brams
And
D. Marc Kilgour

RR# 98-10 **April 1998**

C.V. STARR CENTER FOR APPLIED ECONOMICS

NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, NY 10003-6687

Fallback Bargaining

Steven J. Brams
Department of Politics
New York University
New York, NY 10003
USA

D. Marc Kilgour
Department of Mathematics
Wilfrid Laurier University
Waterloo, Ontario N2L 3C5
CANADA

April 1998

Abstract

Fallback bargaining is a bargaining procedure under which bargainers begin by indicating their preference rankings over all alternatives. They then fall back, in lockstep, to less and less preferred alternatives—starting with first choices, then adding second choices, and so on—until an alternative is found on which all bargainers agree. This common agreement, which becomes the outcome of the procedure, may be different if a decision rule other than unanimity is used. The outcome is always Pareto-optimal but need not be unique; if unanimity is used, it is at least middling in everybody's ranking.

Fallback bargaining may not select a Condorcet alternative, or even the first choice of a majority of bargainers. However, it does maximize bargainers' minimum "satisfaction." When bargainers are allowed to indicate "impasse" in their rankings—below which they would not descend because they prefer no agreement to any lower-level alternative—then impasse itself may become the outcome, foreclosing any agreement.

The vulnerability of fallback bargaining to manipulation is analyzed in terms of both best responses and Nash equilibria. Although a bargainer can sometimes achieve a preferred outcome through an untruthful announcement, the risk of a mutually worst outcome in a Chicken-type game may well deter the bargainers from attempting to be exploitative, especially when information is incomplete.

Fallback bargaining seems useful as a practicable procedure if a set of "reasonable" alternatives can be generated. It leapfrogs the give-and-take of conventional bargaining, which often bogs down in details, by finding a suitable settlement through the simultaneous consideration of all alternatives.

JEL Classification: *D*58, *D*63, *D*71. *Keywords:* Bargaining; impasse; social choice; Condorcet winner; implementation; Nash equilibrium.

Fallback Bargaining[1]

Once your fall-back positions are published, you have already fallen back to them (Eban, 1998, p. 81)

1. Introduction

If two bargainers are in a dispute, at least one must retreat to a fallback position to reach a settlement. "Fallback" in our title would therefore seem redundant—*all* bargaining involves at least one bargainer's falling back to a less preferred position in order to produce a settlement acceptable to both.

To be sure, it would seem unfair, and not "true" bargaining, if one bargainer simply caved in to the demands of the other. Accordingly, we ask the following question: Is there a procedure that facilitates a compromise, whereby bargainers retreat from their most-preferred positions in order to achieve an equitable outcome?

We propose such a procedure, called "fallback bargaining," whereby all bargainers—not necessarily just two—fall back in lockstep to less and less preferred positions until they agree on an outcome. In a variant of this procedure, we allow the lockstep to be broken if a bargainer reaches a point where it prefers no agreement ("impasse") to any alternative that is ranked lower.

Among other things, we demonstrate that fallback bargaining yields an outcome that is Pareto-optimal and, if there is no impasse, at least "middling" for all the bargainers. Moreover, when restricted to just two bargainers, the procedure is difficult (and frequently impossible) to manipulate, even if the bargainers have complete information about each other's preferences.

---

[1]Steven J. Brams gratefully acknowledges the support of the C. V. Starr Center for Applied Economics at New York University. D. Marc Kilgour acknowledges the support of the Social Sciences and Humanities Research Council of Canada and the Laurier Centre for Military Strategic and Disarmament Studies. We thank Jeffrey S. Banks for valuable comments on an earlier draft.

Our model extends a bargaining model of Brams and Doherty (1993; see also Brams, 1994, ch. 7), which presumed that a simple majority, measured by weighted or unweighted votes, must agree on an alternative. In the context of voting, Sertel and Yilmaz (1997) and Hurwicz and Sertel (1997) developed related models, which we shall say more about later, that focus on majority rule. While we focus on unanimity as the decision rule, which seems appropriate in many bargaining situations, we give general results for any decision rule and for any number of bargainers.

The plan of this paper is as follows. In section 2, we describe and illustrate fallback bargaining, assuming a decision rule of unanimity. We determine the maximum "depth" to which the bargainers might have to descend before reaching a common agreement, indicating the parameters to which this depth is sensitive. In addition, we prove that fallback bargaining always leads to a Pareto-optimal outcome that maximizes the minimum satisfaction of all the bargainers (i.e., whose depth for the worst-off bargainer is minimal).[2]

In section 3, we compare the situation in which unanimous consent is required to that in which only a simple or qualified majority of bargainers must agree, as is true in most voting situations. We illustrate, among other things, how an outcome under the

---

[2]Proofs of all theorems, corollaries, and lemmata are given in the Appendix—where their numbering is sometimes different and preceded by *A*'s—but the numbers in the text are always keyed to the Appendix numbers. We follow this unusual convention to

- facilitate exposition of the material in the text, where we start with fallback bargaining with a decision rule of unanimity (section 2) and then generalize to all possible decision rules (section 3);

- save space in the Appendix, where we make some theorems general from the start.

Thus, instead of proving that fallback bargaining maximizes the minimum satisfaction of all bargainers (Theorem 3), we prove in the Appendix that under any decision rule q, fallback bargaining maximizes the minimum satisfaction of at least q bargainers (Theorem $A2$), where q can range from 1 to n.

unanimity rule may be less preferred by a majority of bargainers than another alternative (i.e., the so-called Condorcet winner), which may even be their first choice.

In section 4, we consider the possibility that a bargainer might reach a point in fallback bargaining whereby it would prefer no agreement, or impasse, to agreement. We show that the inclusion of "impasse" in bargainers' preference orders may lead to Pareto-optimal outcomes quite different from those without impasse. These, we suggest, can be observed in real-life bargaining situations in which bargainers, at some point, refuse to compromise, preferring stalemate instead.

In section 5, we restrict attention to two-person bargaining situations, which are by far the most common (Brams, 1990), but we place no restrictions on the number of possible outcomes. If there are only two alternatives (e.g., one bargainer wins, the other loses), it is always optimal for the bargainers to be truthful in ranking alternatives. But truthfulness is not always optimal when there are more than two alternatives, as we demonstrate with two theorems that characterize the best response of one bargainer to the other's truthful ranking.

While we leave a characterization of all Nash equilibria in two-person fallback bargaining games to the Appendix, we offer a quantitative analysis of Nash equilibria in three-outcome and four-outcome games in section 6. Although bargainers can benefit from not being truthful in some games, we argue that they may refrain from trying to be exploitative in others that are vulnerable to misrepresentation—in particular, a class of Chicken-like games—because of the serious risks involved should the other bargainer act similarly. Thus, both bargainers may be deterred from acting strategically.[3]

[3]To be sure, Gibbard (1973), Satterthwaite (1975), and subsequent impossibility results in the social-choice and game-theoretic literature establish that virtually no bargaining or voting procedures are immune from manipulation. These results say little, however, about the kinds of games that may be played, and their specific vulnerabilities, under different procedures. In the case of fallback bargaining, we will argue that, practically speaking, it would be a difficult procedure to manipulate, especially in games of incomplete information.

In section 7, we suggest the kinds of disputes in which bargainers are most likely to benefit from fallback bargaining, or to act as if they use it. These informal uses of the procedure lead us to ask the following questions:

1. What real-life disputes would be most amenable to the formal use of fallback bargaining?
2. Is the resolution of such disputes likely to be fairer, in some sense, than the resolution that would be achieved without the formal procedure?

We conclude that invoking the formal procedure could facilitate the resolution of certain kinds of disputes, but it will require considerable care in the generation of alternatives to which it is applied.

2. Description and Properties

Assume that there are n bargainers, and the set of alternatives (possible agreements) is K, where $|K| = k$. Each bargainer has a strict preference ranking over the k alternatives; all rankings can be represented by an n x k matrix, A, whose (i,j)-entry is a_{ij}. Each ranking is given in descending order: bargainer i's most preferred alternative is a_{i1}, its least preferred a_{ik}.

To illustrate, suppose the set of alternatives is $K = \{a, b, c, d\}$, so $k = 4$. Suppose there are $n = 2$ bargainers, whose preferences are given by

$$A^1 = \begin{pmatrix} a & b & c & d \\ b & d & a & c \end{pmatrix}.$$

Bargainer 1's preference ranking (*abcd*) is indicated by the first row—from a most preferred to d least preferred—and bargainer 2's (*bdac*) by the second row.

Fallback bargaining proceeds as follows:

1. The most-preferred alternative of each bargainer is considered. If this is the same for all bargainers, then this common agreement is the bargaining outcome. The procedure stops, and we call this a depth 1 agreement.

2. If there is no common agreement at depth 1 (i.e., not all the bargainers agree on a most-preferred alternative), then the next-most preferred alternatives of all the bargainers are considered. Any alternative within the top two of *every* bargainer is a depth 2 agreement (there may be either one or two common agreements at depth 2, as we will illustrate shortly). If there is a depth 2 agreement, the procedure stops; otherwise, it continues.

3. As long as there is no common agreement, the bargainers descend—one level at a time—to lower and lower levels in their rankings until the intersection of their top-ranked alternatives becomes, for the first time, nonempty. We call the set of common agreements, when the procedure stops at depth d^*, $CS(A)$, or the *Compromise Set (CS)* of fallback bargaining for matrix A.

Examples. The vertical lines in the following 2-bargainer, 4-alternative examples below indicate the column in the matrix, going from left to right, at which a common agreement first appears. In these four examples, the depth of the agreement, d^*, varies from 1 (A^3) to 3 (A^4). Observe that in two of the examples (A^1 and A^3) the Compromise Sets are singletons, whereas in the other two examples (A^2 and A^4) the Compromise Sets contain two alternatives:

$$A^1 = \begin{pmatrix} a & b & c & d \\ b & d & a & c \end{pmatrix}, \quad CS(A^1) = \{b\}.$$

$$A^2 = \begin{pmatrix} a & b & c & d \\ b & a & c & d \end{pmatrix}, \quad CS(A^2) = \{a, b\}.$$

$$A^3 = \begin{pmatrix} a & b & c & d \\ a & d & b & c \end{pmatrix}, \quad CS(A^3) = \{a\}.$$

$$A^4 = \begin{pmatrix} a & b & c & d \\ c & d & a & b \end{pmatrix}, \quad CS(A^4) = \{a, c\}.$$

What is the maximum depth at which a common agreement can appear? The upper bound is given by

Theorem 1. d^* • |_k - k/n + 1_|.

Proof. See the proof of Theorem *A*1 in the Appendix; an equivalent theorem for the so-called Kant-Rawls Social Compromise, which also assumes unanimity, is given in Hurwicz and Sertel (1997). The proof of the general case—allowing for any decision rule (not just unanimity), which includes Hurwicz and Sertel's (1997) so-called Majoritarian Compromise (simple majority)—is given for Theorem *A*3 in the Appendix.

Among the earlier examples, A^4 shows that the upper bound in Theorem 1 can be attained: d^* = 3 = |_4 - 4/2 + 1_|; in the Appendix, we show that this bound is always tight. We next consider how the upper bound of Theorem 1 behaves as n or as k increases:

1. Dependence on *n*. As the number n of bargainers increases, but the number k of alternatives remains fixed, the upper bound on depth eventually reaches k, the number of alternatives. Thus, if $k = 4$ (as in our previous examples), and n increases from 2 to 4, fallback bargaining may have to descend to $d^* = 4$ before the Compromise Set becomes nonempty, as illustrated by the following example:

$$\mathrm{A}^5 = \begin{pmatrix} a & b & c & d \\ b & c & d & a \\ c & d & a & b \\ d & a & b & c \end{pmatrix}, \qquad CS(A^5) = \{a, b, c, d\}.$$

For any preference matrix A, and any alternative $x \in K$, we define $m(x) = m(x, A)$, the *mean depth of x in A*, to equal the (arithmetic) average rank of x over all rows of A. For instance, $m(a) = m(x, A^5) = (1 + 4 + 3 + 2)/4 = 2.5$.

In A^5, the maximal depth of the agreement is mitigated by the fact that all four alternatives are in the Compromise Set. If one of these is selected at random as the outcome, then the probability that a bargainer will suffer its worst outcome is only 25%.

Moreover, the mean depth of the outcome thus selected is $m(a) = m(b) = m(c) = m(d) =$ 2.5, so on average each bargainer obtains a middling outcome.

2. Dependence on *k*. As the number k of alternatives increases, but the number of bargainers n remains fixed, the upper bound on depth approaches $(1 - 1/n)k + 1 = [(n - 1)/n]k + 1$, or a fixed fraction of k plus 1. For example, if $n = 2$ (as in our first four examples), and k increases from 4 to 8, the upper bound given by Theorem 1 increases from $d^* = 3$ (illustrated by A^4) to $d^* = 5$, which is illustrated by example A^6:

$$A^6 = \begin{pmatrix} a & b & c & d & e & f & g & h \\ e & f & g & h & a & b & c & d \end{pmatrix}, \quad CS(A^6) = \{a, e\}.$$

In this example, the worst-case scenario for each bargainer is a below-average alternative (5th out of 8), whereas the best-case scenario is a best alternative (1st out of 8). Thus, $m(a, A^6) = m(e, A^6) = (1 + 5)/2 = 3$.

As the examples above illustrate, fallback bargaining yields, on average, outcomes that are at least middling for each bargainer, whatever the number of bargainers or the number of alternatives. In the case of two bargainers and two alternatives, a and b—in which one bargainer prefers a to b and the other b to a—the Compromise Set will be simply $\{a, b\}$. This, of course, is hardly a resolution of their bargaining problem.

If, however, both bargainers think some compromise alternative c is better than the other bargainer's preferred alternative, then c will be the outcome of fallback bargaining if it is included as an alternative:

$$A^7 = \begin{pmatrix} a & c & b \\ b & c & a \end{pmatrix}, \quad CS(A^7) = \{c\}.$$

In the concluding section, we will suggest how, through the introduction of *several* compromise alternatives, fallback bargaining can be rendered a useful practical device for finding an acceptable resolution.

Next we give lower and upper bounds on the number of alternatives in the Compromise Set:

Lemma 1. $1 \bullet |\,\mathrm{CS}\,| \bullet \min\{d^*, n\}$

Proof. See the proof of Lemma A1 in the Appendix. The proof of the general case for any decision rule is given for Lemma A2 in the Appendix.

The lower bound on the number of outcomes in the Compromise Set is illustrated by A^1, A^3, and A^7, and the upper bound by A^2, A^4, and A^5.

Of course, if the number of alternatives in the Compromise Set is large, as in example A^5 [$|CS(A^5)| = 4$], there is no ready resolution of the bargaining problem. But this result, we would argue, is to be expected in this example, because majorities cycle: $a > b > c > d > a$, where ">" indicates majority preference. Later we shall consider examples in which preferences are not cyclical and ask whether fallback bargaining chooses a "desirable" alternative.

Next we ask whether fallback bargaining always leads to a Pareto-optimal outcome.[4]

Theorem 2. *If $x \in CS$, then x is Pareto-optimal.*

Proof. See the proof of Theorem A4 in the Appendix, which also covers the general case (i.e., any decision rule).

Example A^1 shows that the converse of Theorem 2 is false: alternative a is Pareto-optimal but does not belong to the Compromise Set. More specifically,

Theorem 3. *The Compromise Set comprises all Pareto-optimal alternatives that maximize the minimum ranking of the bargainers.*

Proof. See the proof of Theorem A2 in the Appendix, which also covers the general case (i.e., any decision rule).

---

[4]Let x and y be any two outcomes. We say x is *Pareto-superior* to y, written $x \succ y$, if all the bargainers rank x higher than y; in this case, y is *Pareto-inferior* to x. If y has the property that there exists no x such that $x \succ y$, then y is *Pareto-optimal*.

Notice in example A^1 that there are two Pareto-optimal alternatives, *a* and *b*. (Alternative *c* is Pareto-inferior to both *a* and *b*, and alternative *d* is Pareto-inferior to *b*.) The lowest ranking that either bargainer assigns to *a* is 3, and the lowest ranking that either bargainer assigns to *b* is 2. Consequently, the Compromise Set is $\{b\}$, which, as shown below, is the first alternative to become common:

- 0 common at depth 1: $\{\varnothing\}$
- 1 common at depth 2: $\{b\}$
- 2 common at depth 3: $\{a, b\}$
- 4 common at depth 4: $\{a, b, c, d\}$.

In examples A^2 - A^6, the Compromise Sets contain all the Pareto-optimal alternatives. By contrast, all three alternatives are Pareto-optimal in example A^7, but only alternative *c* is in the Compromise Set: it becomes common at depth 2, whereas the Pareto-optimal alternatives, *a* and *b*, do not become common until depth 3.

In summary, the Compromise Set produces outcomes that are Pareto-optimal (Theorem 2) and at least middling (Theorem 1), based on their depth or mean depth. These outcomes also maximize the minimum satisfaction that any bargainer enjoys (Theorem 3): the lowest rank given by a bargainer to any alternative not in the Compromise Set, even if it is Pareto-optimal, is always less.[5]

3. Alternative Decision Rules

The choice of a middling outcome may be controversial, as example A^8 illustrates:

---

[5]Fallback bargaining, however, fails one of Arrow's (1963) contions (as it must): independence from irrelevant alternatives. Thus in A^2, if the preference ranking of bargainer 2 changes from *bacd* to *cbad* (i.e., bargainer 2 moves *c* up from third to first place without changing the ranking of the other alternatives), the Compromise Set would change from $\{a, b\}$ to $\{b\}$. In other words, the preference of bargainer 2 for "irrelevant alternative" *c* affects the social choice between *a* and *b*, lowering *a* in the social ordering by singling out *b* as the unique social choice.

$$A^8 = \begin{pmatrix} a & b & c & d \\ a & b & c & d \\ a & d & b & c \\ d & c & b & a \end{pmatrix}, \quad CS(A^8) = \{b\}.$$

Notice that alternative *b* is at rank 2 for two bargainers and at rank 3 for two bargainers, giving it a mean rank of 2.50. By comparison, alternative *a* is at rank 1 for three bargainers and at rank 4 for one bargainer, so its mean rank is 1.75. Moreover, not only is *a* the *Condorcet winner* (in pairwise contests, a majority prefers it to every other alternative), but it is also ranked first by three of the four bargainers. Nevertheless, fallback bargaining chooses *b*, whose only merit seems to be that nobody dislikes it too much (by ranking it last).

The choice of *b*, we believe, is quite indefensible in a voting situation. Indeed, not only would a Condorcet voting procedure select *a*, but so would virtually all other voting procedures, including the Borda count and the Hare system of single transferable vote. On the other hand, insofar as *unanimous consent* is required in a bargaining situation, the choice of *b* seems to us entirely appropriate.

To be sure, if majority support were deemed sufficient to reach a consensus in a bargaining situation, then fallback bargaining could be modified to reflect this less stringent decision rule. With this modification, *a* would be chosen at the outset (i.e., at depth 1) in A^8.[6]

In all 2-person bargaining situations, of course, the unanimity decision rule is the same as the majority decision rule, so there would be no change of outcome sets in our previous 2-person examples. Neither would there be in our earlier 4-person example (A^5), in which majorities cycle (i.e., there is no Condorcet outcome).

---

[6]If the Condorcet winner, *a*, were deleted from every bargainer's ranking, the Compromise Set would expand from {*b*} to {*b*, *c*, *d*}: all three remaining (Pareto-optimal) alternatives would be selected by fallback bargaining. In fact, all four alternatives are Pareto-optimal in A^8; fallback bargaining singles out *b* as the outcome, in contrast to the simple-majority outcome, *a*.

In general, however, the decision rule will make a difference in *n*-person bargaining situations. Consequently, bargainers must decide what is an appropriate level of consensus to require in order to make an agreement binding on all parties.

We next can generalize fallback bargaining to *q-approval fallback bargaining*, in which the decision rule is that, for acceptance, the approval of at least q (for quota) bargainers is required, where q lies between 1 and n inclusive. Until now, we have used the decision rule $q = n$ (i.e., unanimity). Normally, we suppose, the quota will be at least a simple majority of bargainers (i.e., q • |_(n + 1)/2_|), but our results apply for any q in the range 1 • q • n.

The use of q-approval fallback bargaining produces the *q-approval Compromise Set*, CS^q, which is the set of alternatives that are approved of, for the first time as the depth increases, by at least q bargainers. For every q, there will be a d_q*; when $q = n$ (under fallback bargaining), $d^* = d_n^*$.

To illustrate the effects of different quotas, consider again example A^8:

$CS^1(A^8) = \{a, d\}$ at depth $d_1^* = 1$

$CS^2(A^8) = \{a\}$ at depth $d_2^*= 1$

$CS^3(A^8) = \{a\}$ at depth $d_3^* = 1$ (simple majority decision rule)

$CS^4(A^8) = CS(A^8) = \{b\}$ at depth $d_4^* = d^* = 3$ (unanimity decision rule).

Alternative *a* is what Hurwicz and Sertel (1997) call the "Majoritarian Compromise," whereas alternative *b*, our fallback-bargaining outcome, is what they call the "Kant-Rawls Social Compromise."[7] In the Appendix, we give generalizations of Theorem 1, Lemma 1, Theorem 2, and Theorem 3, which we summarize as follows:

---

[7]The Majoritarian Compromise will always be selected from the left half of alternatives in *A* (Sertel and Yilmaz, 1997; Hurwicz and Sertel, 1997), whereas the Kant-Rawls Social Compromise might force a descent to the greatest possible depth (i.e., the right-most column), so one bargainer (or more) might obtain its worst alternative.

General Results for *q*-Level Fallback Bargaining. *For 1 • q • n,*

• *the upper bound on depth is* d_q^* • *|_(kq - k + n)/n_| (Theorem A3);*

• *bounds on the size of* CS^q *are 1 < |*CS^q*| < min{nd*/q, n} (Lemma A2);*

• *if x* $\in$ CS^q*, then x is Pareto-optimal (Theorem A4)*

• *the alternatives in* CS^q *maximize the minimum satisfaction of the q most satisfied bargainers (Theorem A2).*

The Pareto-optimality of all alternatives chosen by *q*-level fallback bargaining, from *q* = 1 to *q* = *n*, is perhaps surprising. This fact, however, is not a good reason for considering all these alternatives to be serious candidates for outcomes of either a bargaining process or a voting process. For example, alternative *d* in example A^8 is in CS^1; but because it is the last choice of two of the four bargainers, it is not an alternative that we would recommend as a consensus choice.

Alternative *c* is the one alternative in example A^8 that is not chosen by *q*-level fallback bargaining for any *q*. Nevertheless, it is Pareto-optimal, demonstrating that the members of all Compromise Sets do not exhaust the set of Pareto-optimal alternatives.

More surprising, perhaps, is that *q* = |_(*n* + 1)/2_| (simple majority), which does choose the Condorcet alternative *a* in example A^8 when fallback bargaining (unanimity) fails to, is not always to so successful, as the following 7-person example by Sertel and Yilmaz (1997) demonstrates:

$$A^9 = \begin{pmatrix} a & b & c & d & e \\ c & b & a & d & e \\ d & b & a & c & e \\ e & b & a & c & d \\ a & d & c & e & b \\ a & e & d & c & b \\ c & d & a & e & b \end{pmatrix}, \quad CS^4 = \{b\}.$$

Alternative *b* is, in fact, the Condorcet loser—majorities prefer it to each of the other alternatives in this example—whereas alternative *a* is the Condorcet winner. Moreover, fallback bargaining (*q* = 7) finds *a* at depth *d** = 3, suggesting it to be a “better” decision rule in this instance (actually, any qualified majority $q > 4$ finds a). Thus, the case for the Majoritarian Compromise, which chooses $\{b\}$, seems weak in example A^9.

It is worth noting that the Condorcet winner may be chosen using non-majority quotas, as the following simple example—in which alternative a is the Condorcet winner—illustrates:

$$A^{10} = \begin{pmatrix} a & b & c \\ a & b & c \\ c & b & a \end{pmatrix}.$$

Notice that a is in the Compromise Sets for both $q = 1$ and $q = 2$, but it is alternative b, chosen by fallback bargaining ($q = 3$), that seems most to deserve the appellation "compromise":

$CS^1(A^{10}) = \{a, c\}$ at depth $d_1^* = 1$

$CS^2(A^{10}) = \{a\}$ at depth $d_2^* = 1$ (simple majority decision rule)

$CS^3(A^{10}) = \{b\}$ at depth $d_3^* = d^* = 2$ (unanimity decision rule).

In example A^{11} below, there is no Condorcet winner because there is a paradox of voting, in which majorities cycle: $c > a > d > e > c$, where ">" indicates the majority preference of the five voters:

$$A^{11} = \begin{pmatrix} a & d & e & c & b \\ a & d & e & c & b \\ b & d & e & c & a \\ b & c & a & d & e \\ c & e & d & a & b \end{pmatrix}.$$

As in example A^{10}, the Compromise Sets in A^{11} yield all possible outcomes:

$CS^1(A^{11}) = \{a, b, c\}$ at depth $d_1^* = 1$

$CS^2(A^{11}) = \{a, b\}$ at depth $d_2^* = 1$

$CS^3(A^{11}) = \{d\}$ at depth $d_3^* = 2$ (simple majority decision rule)

$CS^4(A^{11}) = \{e\}$ at depth $d_4^* = 3$ (qualified majority decision rule)

$CS^5(A^{11}) = CS(A^{11}) = \{c\}$ at depth $d_5^* = d^* = 4$ (unanimity).

Not only is alternative *d*, the simple-majority outcome, not a Condorcet winner, but $q = 4$ (qualified majority) gives a different outcome from either $q = 3$ (simple majority) or $q = 5$ (unanimity). Because all five alternatives are chosen by the various *q*'s, a consensus choice is by no means evident.

Clearly, the decision rule, even when restricted to a simple majority or greater, can make a big difference in the outcome. In the concluding section, we will turn to the question of what level of consensus *should* be required in order to implement a compromise agreement.

4. The Effects of Impasse

We next consider the possibility that bargainers set limits—or "reservation prices," in the vernacular of economics—on how low they will dip in their rankings before "throwing in the towel" (i.e., giving up rather than accepting a less-preferred agreement). Specifically, assume that each bargainer puts *I* (for impasse) in its preference ranking at the level at which it prefers no agreement to any lower-level alternative. We call this modification of fallback bargaining *fallback bargaining with impasse*.[8]

Fallback bargaining with impasse proceeds exactly as does fallback bargaining, but with one restriction. Once the descent process reaches *I* for a bargainer, it stops for that bargainer. If no common agreement is reached by the time the level descends to every bargainer's *I*, *I*—not an alternative in *K*—is the outcome.

In fallback bargaining with impasse, the Compromise Set is called *CSI*, and it is reached at depth d^*. If there are *n* bargainers and the alternative set is *K*, where $|K| = k$,

---

[8]Brams and Doherty (1993; see also Brams, 1994, ch. 7) were the first to introduce impasse into the preferences of bargainers. Like Sertel and Yilmaz (1997) and Hurwicz and Sertel (1997), they assumed that only a simple majority of bargainers need agree on an alternative in order for it to be chosen. Unlike the present model, however, *I* can be breached in the Brams-Doherty (1993) model: a bargainer will support an alternative below *I* if there is another alternative that would otherwise be chosen that the bargainer ranks still lower than *I*. Thus, *I* is not an impregnable barrier in their model.

then preferences are given not by an n x k matrix A but by an n x $(k + 1)$ matrix B, in which the i^{th} row gives bargainer i's ranking of $K \cup \{I\}$.

Examples. In the following four examples, the preference rankings of the two bargainers for the set of four possible agreements, $\{a, b, c, d\}$, duplicate those of example A^4 earlier. Now, however, the appearance of I at different levels in each bargainer's ranking may change the Compromise Sets from $CS(A^4) = \{a, c\}$:[9]

$$B^1 = \begin{pmatrix} a & I & b & c & d \\ c & I & d & a & b \end{pmatrix}, \quad CSI(B^1) = \{I\}$$

$$B^2 = \begin{pmatrix} a & I & b & c & d \\ c & d & a & I & b \end{pmatrix}, \quad CSI(B^2) = \{a\}$$

$$B^3 = \begin{pmatrix} a & b & c & I & d \\ c & d & I & a & b \end{pmatrix}, \quad CSI(B^3) = \{c\}$$

$$B^4 = \begin{pmatrix} a & b & c & I & d \\ c & d & a & b & I \end{pmatrix}, \quad CSI(B^4) = \{a, c\}.$$

Notice that the common agreements in the Compromise Sets are either Pareto-superior to I or I itself. Moreover, if either a or c is Pareto-superior to I, then it falls in the Compromise Set. That this is no accident is shown by

Theorem 4. *Let the n x (k + 1) matrix B describe the preferences of the bargainers under fallback bargaining with impasse. Let $x \in K$, the set of alternatives, and construct the n x k matrix A by deleting I from each row of B. Then $x \in CSI(B)$ if $x \succ I$ (i.e., x is Pareto-superior to I) and $x \in CS(A)$; if $I \in CSI(B)$, then I is the unique member of CSI(B).*

Proof. See Theorem $A5$ in the Appendix.

Corollary 1. *All elements of CSI(B) are Pareto-optimal.*

---

[9]As before, in each example we indicate with vertical lines the level, going from left to right, at which common agreement(s) or I first appear. Notice that in all the examples except B^4, at least one bargainer "reaches" I, but it never descends past it.

Proof. See Corollary *A*1 in the Appendix.

Let $x \in K$. Theorem 4 states that if $x \in CS(A)$ and $x \succ I$, then $x \in CSI(B)$. The converse, however, is not true. If $x \in CS(B)$, then $x \succ I$, but it is nonetheless possible that $x \notin CS(A)$, as shown by the following example:

$$B^5 = \begin{pmatrix} a & I & b & c \\ c & b & a & I \end{pmatrix}, \quad CSI(B^5) = \{a\}; \qquad A^5 = \begin{pmatrix} a & b & c \\ c & b & a \end{pmatrix}, \quad CS(A^5) = \{b\}.$$

Here *a*, which is Pareto-superior to *I*, is the only member of the Compromise Set with impasse, whereas *b* is the only member of the corresponding Compromise Set without impasse.

It follows from Theorem 4 that *I* is the unique member of the Compromise Set with impasse iff *I* is Pareto-optimal. To determine *CSI*, one can begin by finding alternatives Pareto-superior to *I*; if there are none, then *CSI* contains only *I*.

An example that illustrates this rule, wherein *b* is the only alternative Pareto-superior to *I*, is

$$B^6 = \begin{pmatrix} a & b & I & c \\ b & I & c & a \\ c & a & b & I \end{pmatrix}, \quad CSI(B^6) = \{b\}.$$

Not surprisingly, the most intransigent bargainer—the one that ranks *I* highest (bargainer 2, or *B*2, whose preferences are given in the second row of B^6)—is the one that gets its most-preferred alternative (*b*).[10]

The choice of *b* in this example might be contested on the ground that a majority of bargainers (*B*1 and *B*3) prefer *a* to *b*, which we write as $a > b$. But, in fact, the story is

[10]If either *B*1 or *B*3 tried to be more intransigent for strategic reasons (i.e., by ranking, against its true preferences, *I* higher than *b* rather than vice versa), then *I* would be the outcome. Thus, being "strategically" intransigent may succeed only in sabotaging a preferred agreement, especially if the other bargainers are acting similarly.

more complicated than this, because majorities cycle: $a > b > I > c > a$, where ">" indicates majority preference. Thus, there is a paradox of voting that includes *I*.[11]

If we rule out *c* on the ground that a majority of bargainers (*B*1 and *B*2) prefer *I* to *c*, this still leaves *a* and *b* as viable possibilities. While *b* is the fallback bargaining outcome and also has the greater mean depth, one wonders whether the fact that $a > b$ should not swing the bargaining choice toward *a*.

In the normative social-choice and voting literature, questions such as this are addressed, but we will not pursue them further here. Suffice it to say that fallback bargaining with impasse produces a set of Pareto-optimal alternatives, or impasse, that maximize the minimum satisfaction of bargainers—but now with *I* excluding certain alternatives that, without *I*, might have been considered satisfactory.

We turn next to an analysis of the vulnerability of fallback bargaining to strategic manipulation, first by characterizing best responses (section 5) and then Nash equilibria (section 6). To keep matters simple, we will restrict the analysis to two bargainers and assume that they cannot indicate impasse. In section 7 we will offer some thoughts on generalizing our results, allowing both for more bargainers and for the possibility of impasse.

5. Vulnerability of Two-Person Fallback Bargaining: Best Responses

In section 2 we discussed the situation in which there are only two alternatives, *a* and *b*, wherein *B*1 prefers *a* to *b* and *B*2 prefers *b* to *a*. If each bargainer truthfully indicates its preference, the Compromise Set is $\{a, b\}$.

---

[11]If *I* were deleted from the rankings of B^6, then the paradox of voting would remain, whereby $a > b > c > a$. In this case, fallback bargaining (without impasse) would give $\{a, b, c\}$ as the Compromise Set, whereas B^6 (with the *I*'s included) singles out *b*. We see nothing wrong with the fact that inclusion of the *I*'s narrows down the outcomes in the Compromise Set.

Neither bargainer can do better by being untruthful. For example, if $B2$ indicated that it, like $B1$, preferred a to b, then it would succeed only in ensuring its less preferred alternative, a.

But with as few as three alternatives, fallback bargaining becomes vulnerable to strategic misrepresentation. To illustrate, assume the truthful preferences of two bargainers are those of example A^{12} below, which gives two outcomes in the Compromise Set:

$$A^{12} = \begin{pmatrix} a & b & c \\ b & a & c \end{pmatrix}, \quad CS(A^{12}) = \{a, b\}.$$

Now if $B1$ announced its preferences to be those shown in the first row of either examples A^{13} or A^{14} below, and $B2$ stuck with its true preferences in the second row, the Compromise Set would be a singleton, containing $B1$'s most-preferred alternative:

$$A^{13} = \begin{pmatrix} a & c & b \\ b & a & c \end{pmatrix}, \quad CS(A^{13}) = \{a\}$$

$$A^{14} = \begin{pmatrix} c & a & b \\ b & a & c \end{pmatrix}, \quad CS(A^{14}) = \{a\}.$$

Thus, $B1$ would have good reason to falsify its preference ranking if it knew (i) $B2$'s true preference ordering and (ii) that $B2$ did not know it was being manipulated in this way (and would have no reason, therefore, not to be truthful).

The relationship of the orderings selected by the bargainers to the resulting Compromise Set constitute the game-form of fallback bargaining (Hurwicz, 1996). For the case of $k = 3$ alternatives, the game-form is shown in Figure 1 (ignore the asteriks for

Figure 1 about here

for now). Note that the game-form does not describe preferences; nonetheless, it is a natural tool to study the consequences of a bargainer's reporting its preferences, either truthfully or untruthfully.

For instance, suppose that *B*1's true preference ordering is *abc*. From Figure 1, one can determine whether reporting *abc*, or some other ordering, gives *B*1 a better outcome. As we show in the Appendix, *B*1 cannot do better than choose *abc* if *B*2's ordering is any of *abc*, *acb*, *bca*, *cba*, or *cab*, but if *B*2's odering is *bac*, *B*1 is better off choosing *acb*.

Whether bargainers can benefit from misrepresenting their preferences rather than reporting them truthfully is essentially the question that Sertel and Yilmaz (1997) and Hurwicz and Sertel (1997) ask concerning the |_(*n* + 1)/2_|-approval Compromise Set (or the Majoritarian Compromise). Like them, we show that Nash-equilibrium implementability is impossible[12]—not for the Majoritarian Compromise, however, but for fallback bargaining. As a prelude to characterizing Nash equilibria under fallback bargaining, we examine the optimal response of one bargainer (truthful or untruthful) to the other bargainer's ranking.

Theorems 5 and 6 below cover, respectively, the cases of an odd and an even number of alternatives *k*:

Theorem 5. *If k = 2h* - 1 *is odd (so h is integral), and B2's ranking, b, is fixed and known to B1, then the best Compromise Set that B1 can achieve is a singleton containing B1's most-preferred alternative among* $b_1, b_2, \ldots, b_h$ *(i.e., among the top h items in B2's ranking). Call this best alternative* b_r*. To achieve* $\{b_r\}$*, B1 submits its true ordering, a, unless* b_r *is ranked (strictly) lower by B2 than by B1, and there are alternatives ranked at or above level r by B1 that are preferred to* b_r *by B2. In this case, B1 can achieve* $\{b_r\}$

---

[12]Sertel and Yilmaz (1997) demonstrate, additionally, that the Majoritarian Compromise is subgame-perfect implementable, but they do not find any "natural" mechanism for effecting such implementation. In the absence of a simple mechanism—whose message space can easily be explained to voters—we will concentrate in section 6 on the conditions under which, when the message space is the bargainers' direct statement of their preferences, truthful revelation is a Nash equilibrium. (Our theorems in section 6, however, characterize all Nash equilbria—both those involving truthful revelation and those involving misrepresenation.)

by switching these alternatives with alternatives that would otherwise be below level r in both B1's and B2's orderings.

Proof. See the proof of Theorem *A*6 in the Appendix.

Examples A^{15}, A^{16}, and A^{17} illustrate Theorem 5.

Examples. Let $k = 7$ and suppose the preference ranking of *B*1 is *abcdefg*. If *B*2's ranking is *edbgacf*, we have, if both *B*1 and *B*2 are truthful,

$$A^{15} = \begin{pmatrix} a & b & c & d & e & f & g \\ e & d & b & c & a & g & f \end{pmatrix}, \quad CS(A^{15}) = \{b\}.$$

Because *B*1's most-preferred alternative up to level $h = 4$ in *B*2's ordering is *b*, the best Compromise Set that *B*1 can achieve is $\{b\}$. With respect to Theorem 5, $r = 3$. While alternative *b* is ranked higher by *B*1 than by *B*2, there are no alternatives ranked at or above level $r = 3$ by *B*1 and preferred to *b* by *B*2. Therefore, *B*1's best response is to be truthful.

Now assume that *B*2's ranking is *cegbadf*. If *B*1 is truthful, we have

$$A^{16} = \begin{pmatrix} a & b & c & d & e & f & g \\ c & e & g & b & a & d & f \end{pmatrix}, \quad CS(A^{16}) = \{c\}.$$

Because *B*1's most-preferred alternative up to level $h = 4$ in *B*2's ordering is *b*, again the best Compromise Set that *B*1 can achieve is $\{b\}$. With respect to Theorem 5, $r = 4$. But because alternative *c* is ranked at or above level 4 by *B*1, and it is preferred to alternative *b* by *B*2, *B*1 can do better by falsifying its preferences. Following Theorem 5, *B*1 can switch alternative *c* with an alternative that is ranked below level 4 in both bargainers' rankings, which is alternative *f* in example A^{16}.

If *B*1 interchanges *c* and *f* in its truthful ranking—announcing *abfdecg* instead—the Compromise Set is indeed $\{b\}$ rather than $\{c\}$:

$$A^{17} = \begin{pmatrix} a & b & f & d & e & c & g \\ c & e & g & b & a & d & f \end{pmatrix}, \quad CS(A^{17}) = \{b\}.$$

Thus, *B*1 obtains its second rather than its third choice when it acts strategically, according to Theorem 5.

Theorem 6. *If $k = 2h$ is even (so h is integral), and B2's ranking, b, is fixed and known to B1, then the best Compromise Set that B1 can achieve is either $\{b_r\}$, containing B1's most-preferred alternative among the top h alternatives in B2's ranking, or $\{b_r, b_{h+1}\}$,. The latter is preferred to the former iff $b_{h+1} > b_r$ (i.e., if b_{h+1} is preferred to b_r according to B1's true ordering, a). To achieve a Compromise Set of the form $\{b_r\}$, B1's choice of ordering is essentially the same as in Theorem 5. To achieve a Compromise Set of the form $\{b_r, b_{h+1}\}$, B1's ordering must place b_r at level $h + 1$, which is a necessary condition for B1 to be able to submit its true ranking. If this condition is not satisfied by B1's truthful ranking, it can always be arranged to do so.*

Proof. See the proof of Theorem *A*7 in the Appendix.

To show the interesting twist that can occur when the number of alternatives is even, let $k = 6$. Suppose that *B*1's true ranking is *abcdef*, and that *B*1 knows that *B*2 will submit the ranking *fcdbea*. If *B*1 submits its true ranking, the Compromise Set will be $\{c\}$. But by reporting its ranking to be *abecdf*, *B*1 can achieve a Compromise Set that it prefers to $\{c\}$:

$$A^{18} = \begin{pmatrix} a & b & e & c & d & f \\ f & c & d & b & e & a \end{pmatrix}, \quad CS(A^{18}) = \{b, c\}.$$

An essential step in constructing *B*1's reported ranking is to move *d* below the fourth level.

Finally, in the Appendix we note that there is no limit to the number of ordinal rankings by which *B*1 can improve the Compromise Set by misrepresenting its preferences. To illustrate this result when there are $k = 9$ alternatives (so $h = 5$), suppose that *B*1's true preference ordering is *abdcefghi*, and that *B*2 submits the ranking *defgabchi*. By reporting its true ranking, *B*1 obtains the Compromise Set {*d*}; but by reporting ranking *abchifgde*, B1 improves the Compromise Set by h - 2 = 3 ordinals:

$$A^{19} = \begin{pmatrix} a & b & c & h & i & f & g & d & e \\ d & e & f & g & a & b & c & h & i \end{pmatrix}, \quad CS(A^{19}) = \{a\}.$$

As shown in the Appendix, this example can be generalized to any value of h (and k = 2h - 1) to show that *B*1 can improve the Compromise Set, relative to truthful reporting, by h - 2 ordinals.

So far we have assumed that *B*1 has complete information about *B*2's ranking of alternatives. Moreover, *B*1 knows what ranking *B*2 will submit (truthful or not), perhaps by having a spy in *B*2's camp. Thereby *B*1 can formulate a best response, which may involve making a false announcement of its preference ranking.

In most bargaining situations, however, it is unlikely that there will be such an asymmetry of information that would allow *B*1 to exploit *B*2 in this manner. Thus, we next turn to an analysis of the game that two bargainers play when they both know each other's preference rankings and must, independently, choose announcement strategies.

6. Vulnerability of Two-Person Fallback Bargaining: Nash Equilibria

To illustrate our general results that characterize all Nash equilibria in two-person fallback bargaining games, we start with the case of k = 3 alternatives, {*a*, *b*, *c*}. Consider the 2 x 2 game between *B*1 with preferences *abc* and *B*2 with preferences *bac*. Assume each bargainer may be either truthful in its announcement (first strategy) or untruthful (second strategy):

		*B*2	
		bac	*bca*
*B*1	*abc*	{*a*, *b*}	<u>{*b*}</u>
	acb	<u>{*a*}</u>	{*c*}

Clearly, each bargainer does better by being untruthful when the other bargainer is truthful—obtaining its best possible Compromise Set ({*a*} for *B*1 and {*b*} for *B*2)—whereas each bargainer gets its worst Compromise Set, {*c*}, when both bargainers are untruthful. In between, both bargainers obtain a middling Compromise Set, {*a*, *b*}, when both are truthful. This game is, in fact, Chicken, and outcomes {*a*} and {*b*}, which we have underscored, represent the two pure-strategy Nash equilibria.[13]

Theorems *A*8 and *A*9 in the Appendix establish necessary and sufficient conditions for pairs of rankings to be Nash equilibria. These theorems, which characterize all Nash equilibria in two-person fallback bargaining games—in which unanimity and majority rule are the same—are analogous to Theorems 5 and 6 in section 6 (there is both an odd and even case).[14]

To continue our development of the $k = 3$ case, the two pure-strategy Nash equilibria shown in the 2 x 2 game become four pure-strategy Nash equilibria in an expanded 3 x 3 game that includes additional strategies *cab* for *B*1 and *cba* for *B*2:

		*B*2		
		bac	*bca*	*cba*
	abc	{*a*, *b*}	<u>{*b*}</u>	<u>{*b*}</u>
*B*1	*acb*	<u>{*a*}</u>	{*c*}	{*c*}
	cab	<u>{*a*}</u>	{*c*}	{*c*}

[13]If the payoffs to the bargainers were in cardinal utilities, there would be a third mixed-strategy Nash equilibrium. Even if we were to assume payoffs in utiles, the mixed-strategy equilibrium would not be compelling because it is Pareto-inferior to {*a*, *b*}.

[14]Because Theorems *A*7 and *A*8 require considerable technical development to state precisely, we thought it more instructive not to repeat this development in the text but, instead, to provide a detailed quantitative analysis of the cases of $k = 3$ and $k = 4$ alternatives. These two cases offer new insights into the kinds of games that two bargainers might play, which the characterization of Nash equilibria given by Theorems *A*7 and *A*8 does not illuminate. These theorems, nevertheless, are used to determine the equilibria in the two cases.

The strategies associated with the four underscored outcomes—all involving truthfulness by one bargainer and untruthfulness by the other—occur in essentially the only two-bargainer, three-alternative cases in which truthfulness on the part of one bargainer is not a best response to truthfulness on the part of the other. In effect, this game is an expanded version of the classic 2 x 2 Chicken game.

But each bargainer can submit any of six possible orderings, producing a more complex, but still Chicken-like, game. Assume that *B*1's preference order is *abc*, and *B*2's is *bac*. The the ten possible pure-strategy Nash equilibria are indicated by asterisks in Figure 1. Observe that if

- *B*1 submits an ordering consistent with any equilibrium {*a*} (its most-preferred outcome); and
- *B*2 submits an ordering consistent with any equilibrium {*b*} (its most-preferred outcome),

the Compromise Set is {*c*}, which is the worst outcome for both bargainers. We repeat, however, that this case (*abc* versus *bac*) is unusual in the sense that if there are three alternatives, and the bargainers' true preference orders are selected equiprobabily, then the probability that truthful reporting is not a Nash equilibrium is only 1/6.

Does the relative invulnerability of fallback bargaining to manipulation hold when there are more than three alternatives? We next consider the case of k = 4 alternatives, {*a*, *b*, *c*, *d*}, which each bargainer can rank in 4! = 24 different ways. Holding fixed *B*1's preference ranking of *abcd*, we have analyzed when being truthful is a Nash equilibrium in all four-alternative situations. Our results follow:

1. Truthfulness. Truthfulness on the part of both bargainers is a Nash equilibrium in 15 of the 24 cases (62.5%). *B*2 can do no better than be truthful if its rankings are as follows:

- *abcd*, *abdc*, *acbd*, *acdb*, *adbc*, and *adcb*, all giving {*a*}
- *bcda* and *bdca*, both giving {*b*};
- *cdab* and *cdba*, giving {*a*, *c*} and {*b*, *c*}, respectively;
- *dabc* and *dacb*, both giving {*a*};
- *dbca*, *dcab*, *dcba*, giving {*b*}, {*a*, *c*}, and {*b*, *c*}, respectively.

2. Chicken. There are two pure-strategy Nash equilibria in the three games in which *B*2's truthful ranking is *bacd*, *badc*, or *cbad* (12.5%). To illustrate these three games, consider the game in which *B*2's truthful ranking is *bacd*. The Compromise Sets associated with Nash equilibria in the following 2 x 2 game, which are underscored, involve one bargainer's being truthful while the other is not:

		*B*2	
		bacd	*bdca*
*B*1	*abcd*	{*a*, *b*}	<u>{*b*}</u>
	adcb	<u>{*a*}</u>	{*d*}

This game is analogous to our earlier 2 x 2 game of Chicken—in which each bargainer ranked only three alternatives—except now the "disastrous" outcome when both bargainers are untruthful (i.e., {*d*}) is ranked fourth rather than third by both of them.

Of course, this 2 x 2 game is only a microcosm of the 24 x 24 game that the bargainers would actually play. Application of Theorem *A*6 shows that several of the other 22 strategies for each bargainer are in equilibrium. For example, *B*1's choice of *acdb*, *acbd*, or *adbc* also results in {*a*} if *B*2 is truthful by announcing *bacd*. Similarly, *B*2 has three additional Nash equilibrium strategies that it can use against *B*1's truthful announcement of *abcd*. All these games are, in a sense, expanded versions of Chicken.

3. Best Response. One bargainer has a best response, which is untruthful, to the other bargainer's truthful announcement, but not vice versa, in 6 of the 24 cases (25%). When such one-sided manipulation is possible, the resulting Nash equilibrium yields a pair of alternatives that favors the untruthful bargainer (1st and 2nd choices) over the truthful bargainer (1st and 3rd choices):

• Assume *B*1 is truthful and announces *abcd*. Then if *B*2's truthful rankings are *cabd*, *cadb*, or *cbad*, *B*2's best responses are *cdab*, *cdab*, and *cdba*, respectively, resulting in {*a*, *c*}, {*a*, *c*}, and {*b*, *c*}.

• Assume *B*2 is truthful and announces *bcad*, *bdac*, or *dbac*. Then *B*1's best responses are *adbc*, *acbd*, and *acbd*, respectively, resulting in {*a*, *b*} in each case.

Discussion. In the games in which *B*1 and *B*2 do not agree on a first choice, the cases vulnerable to strategic manipulation increase from 16.7% (1 out of 6) in the three-alternative case to 37.5% (9 out of 24) in the four-alternative case. We hypothesize that the strategic incentives for manipulation continue to increase, even on a proportional basis, as the number of alternatives increases.

On the other hand, the benefits of manipulation may be illusory in some of these situations. Because the game is Chicken in the one case vulnerable to manipulation when $k = 3$, and in three cases when $k = 4$, the risks are as great as the benefits. Instead of ensuring at least a next-best outcome by being truthful, a bargainer choosing an optimal manipulative strategy in Chicken risks its worst outcome in attempting to obtain its best outcome. We believe that many bargainers, using fallback bargaining, would choose *not* to court disaster in Chicken, even if they had complete information about their opponent's preferences.

By contrast, the six best-response cases in the four-outcome games are "safe" in the sense that the other bargainer has no counter-response that yields it a better outcome. Hence, the untruthful bargainer can use a manipulative strategy with impunity, because a rational opponent has no recourse—it can do no better than be truthful.

But the benefits of such manipulation are not great: an optimal strategy in the best-response cases ensures the untruthful player of either its best or next-best outcome instead of, unfailingly, its next-best outcome when it is truthful.[15] Furthermore, because *each* bargainer has the opportunity of using a best-response strategy in only three of the 24 cases (12.5%), there are relatively few occasions in which to exploit such a strategy.

In the best-response cases, notice that the preferences of the bargainers are neither coincident nor diametrically opposed. Thus, for example, if *B*1 truthfully announces *abcd*, *B*2 can be exploitative by being untruthful only if its truthful preferences are *cabd*, *cadb*, or *cbad*. In each case, what is best for *B*2 is next-worst for *B*1.

Most real-life bargaining situations, of course, are suffused with incomplete information, to which our theoretical results on Nash equilibria in the three-outcome and four-outcome complete-information games are not applicable. Nevertheless, there is certainly something Chicken-like in many bargaining situations; our findings support this view, suggesting risks even when information is complete.

Under fallback bargaining, the bargainers' rankings of alternatives completely determine which one(s) will be chosen. Because a bargainer has less opportunity to "feel out" the resolve of an opponent when this procedure is invoked than when there is endless haggling, we think that fallback bargainers will be reluctant to risk disaster, at least in Chicken. Moreover, if the bargainers have incomplete information about each other's preferences, they may not even know that they are playing Chicken—or one of the more numerous games in which truthfulness is an optimal strategy. Thus, it will often pay for them to exercise caution by being truthful under fallback bargaining.

7. Conclusions

Possible Bias. While fallback bargaining seems a promising procedure for inducing compromises, it is legitimate to ask how the alternatives on which the procedure

---

[15]To be sure, a reasonable chance of getting one's best alternative, versus negotiating a compromise that reflects an "average" of one's best and next-best alternatives, will not be trivial if there is a big difference between these two alternatives, as measured by their cardinal-utility values.

operates might be generated. If the alternatives are strongly biased in favor of one bargainer, then the procedure's selection of a Pareto-optimal and middling outcome is a cruel joke against the bargainer or bargainers suffering this bias.

Generating Alternatives. One possible solution to the bias problem is to allow the bargainers themselves to propose different alternatives. If the bargaining is over a future contract between labor and management, for example, each side could propose alternative agreements that have similar cost implications, albeit in opposite directions. A neutral party might be used to assess that each side's alternative proposals more or less match the gains and losses of each other, or are equidistant from the status quo. Fallback bargaining would then enable the bargainers to leapfrog the give-and-take of conventional bargaining, which often bogs down in details, by finding a suitable settlement through the simultaneous consideration of all alternatives.

Correspondence to Real-Life Compromises. We think the give-and-take of conventional bargaining, especially that which results in the successful settlements of disputes, often approximates what fallback bargaining formalizes. This is probably especially true in business disputes, in which the costs and benefits of alternative agreements can often be calculated with some precision.

In personal disputes, including divorce, this assessment is undoubtedly harder. For this purpose procedures like "adjusted winner," in which the parties can allocate points over the items (goods or issues) in dispute, may be more practicable (Brams and Taylor, 1996).

Normative Advantages. We believe that one important advantage of fallback bargaining over the give-and-take of traditional negotiations is that it allows for many different proposals to be on the table at once. Moreover, no decisions need be made about features that are acceptable or not acceptable—as in the usual step-by-step bargaining—in order for the parties to "advance" to a next stage. Indeed, advancing in fallback bargaining means generating new ideas, or packaging them in different ways, both of which may be facilitated by a mediator's making independent proposals, or combining parts of old ones, that are then put on the table as new alternatives.

In the end, of course, these alternatives must be ranked by the parties, but it may not be clear, even to them, how best to do this until there are no more proposals put forward. Although it is possible that the parties will try to anticipate each others' choices and strategize when they announce their rankings, we think this would be extremely difficult if there are, say, ten or more alternatives on the table.[16] As we have illustrated with only three or four alternatives and two bargainers, misrepresenting one's preferences may sometimes allow one to reap large benefits—but can dangerously backfire as well, to the detriment of all.

The Role of Impasse. We have mixed feelings about allowing the bargainers to incorporate impasse (*I*) into their rankings. On the one hand, some bargainers might genuinely prefer *I* to any agreement below it and should be entitled to so indicate their position. On the other hand, the inclusion of *I* may lead to continuing disagreement that is not authentic, because it allows for, and may even encourage, strategizing beyond that which would occur without *I*.

Without *I*, fallback bargaining ensures some agreement and, therefore, closure of the negotiation process. But the middling outcome produced may not be a Condorcet winner, as we showed, or possess other desirable features.

Fairness. If there are more than two bargainers, it may be advisable to relax the unanimity rule, assuming this is acceptable to the bargainers. A simple or qualified-majority decision rule will, in general, be more likely to find Condorcet alternatives, if

---

[16]When there are relatively few alternatives, fallback bargaining is more likely to produce ties, which suggests that there needs to be a device for selecting one alternative from Compromise Sets that contain more than one. Sertel and Yilmaz (1997) offer different suggestions for breaking ties, including choosing the alternative with the most first-choice approvals, then second-choice approvals, and so on. One might also choose the alternative with the highest average approval.

they exist, than does fallback bargaining (with unanimity).[17] But the choice of a Condorcet alternative could be at the price of inflicting on some bargainers outcomes that, while Pareto-optimal, are quite damaging. In such cases, a middling outcome for everybody may be fairer.

We mention these effects of the inclusion of *I*, and the weakening of the unanimity decision rule, to underscore that, as always, there are trade-offs. Although most of our analysis has been of fallback bargaining, which assumed unanimity and was often for only two bargainers, further consideration of the multibargainer case, with a non-unanimous decision rule and the inclusion of *I*, would be desirable.[18]

---

[17]On these and other grounds, Sertel and Sanver (1997) argue that the Majoritarian Compromise, which presumes a simple-majority decision rule, would be desirable in certain kinds of elections in Turkey. Also at a normative level, Brams and Fishburn (1983) argue for approval voting in multicandidate elections. The approval-voting winner and the Majoritarian Compromise will coincide if (i) all voters indicate the same level of approval in making their selections and (ii) that level is the depth at which a majority winner appears for the first time. Approval voting, in contrast to *q*-approval fallback bargaining, makes no presumption of the lockstep descent of all voters to lower and lower levels of approval until a *q*-approval winner appears; instead, it leaves open to the voter where he or she draws the line between acceptable and unacceptable candidates. Although this is somewhat akin to bargainers' indicating *I* in their preference rankings—below which alternatives are unacceptable—there is no descent process under approval voting: all candidates approved of receive their votes at the start. Insofar as voters do rank candidates, their rankings are unexpressed and, therefore, invidious.

[18]In the context of elections, it is worth mentioning those for a council or legislature, in which there are multiple winners and one wishes to achieve proportional representation (PR). In this situation, we believe it sensible to set *q* relatively low (e.g., the size of the electorate divided by the size of the legislature, or the size of an average constituency). Then one would elect candidates who are ranked relatively high by relatively few voters—the number in an average-size constituency—which would help to ensure that a diversity of views is represented and thereby satisfy PR. Related PR systems are analyzed in Potthoff and Brams (1998).

References

Arrow, Kenneth J. (1963). *Social Choice and Individual Values,* 2d ed. New Haven, CT: Yale University Press.

Brams, Steven J. (1990). *Negotiation Games: Applying Game Theory to Bargaining and Arbitration*. New York: Routledge.

Brams, Steven J. (1994). *Theory of Moves*. Cambridge, UK: Cambridge University Press.

Brams, Steven J., and Ann E. Doherty (1993). "Intransigence in Negotiations: The Dynamics of Disagreement." *Journal of Conflict Resolution* 37, no. 4 (December): 692-708.

Brams, Steven J., and Peter C. Fishburn (1983). *Approval Voting*. Boston: Birkhäuser/Springer-Verlag.

Brams, Steven J., and Alan D. Taylor (1996). *Fair Division: From Cake-Cutting to Dispute Resolution*. Cambridge, UK: Cambridge University Press.

Eban, Abba (1998). *Diplomacy for the Next Century*. New Haven, CT: Yale University Press.

Gibbard, Alan (1973). "Manipulation of Voting Schemes: A General Result." *Econometrica* 41 (May): 587-601.

Hurwicz, Leonid (1996). "Institutions as Families of Game Forms." *Japan Economic Review* 27, no. 2: 113-132.

Hurwicz, Leonid, and Murat R. Sertel (1997). "Designing Mechanisms, in Particular for Electoral Systems: The Majoritarian Compromise." Preprint, Department of Economics, Bogazici University, Istanbul, Turkey.

Potthoff, Richard F., and Steven J. Brams (1998). "Proportional Representation: Broadening the Options." *Journal of Theoretical Politics*, forthcoming.

Satterthwaite, Mark Alllen (1975). “Strategy Proofness and Arrow’s Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions.” *Journal of Economic Theory* 10 (April): 197-218.

Sertel, Murat R., and M. Remzi Sanver (1997). “Designing Public Choice Mechanisms: Alternative Electoral Systems for Turkey.” Preprint, Department of Economics, Bogazici University, Istanbul, Turkey.

Sertel, Murat R., and Bilge Yilmaz (1997) “The Majoritarian Compromise Is Majoritarian-Optimal and Subgame-Perfect Implementable.” Preprint, Department of Economics, Bogazici University, Istanbul, Turkey.

APPENDIX

Assume there are n bargainers, and the set of alternatives (possible agreements) is K, where $|K| = k$. Preferences of the bargainers over the alternatives are specified by an $n \times k$ matrix $A = (a_{ij})$ such that, for $i = 1, 2, \ldots, n$, the i^{th} row of A is a permutation of K representing the ranking (in descending order) of bargainer i.

For each bargainer, i, and each alternative $x \in K$, define $j_i^A(x)$ to be the value of j satisfying $a_{ij} = x$. Thus, $j_i^A(x)$ is bargainer i's ranking for alternative x. Define i's satisfaction with alternative x by $h_i^A(x) = k - j_i^A(x)$. Usually, $j_i^A(x)$ and $h_i^A(x)$ can be denoted $j_i(x)$ and $h_i(x)$, respectively, with no possibility of confusion. Note that a_{i1} is i's most preferred alternative, so $j_i^A(a_{i1}) = 1$ and $h_i^A(a_{i1}) = k - 1$.

To describe fallback bargaining formally, define the set of *depth d agreements* as

$$CS_d = \left\{ x \in K : j_i(x) \le d \quad \forall\ i = 1, 2, \ldots, n \right\}$$

for $d = 0, 1, 2, \ldots, k$. Note that CS_d is the set of all alternatives that are among the top d in the ranking of every bargainer. Thus, any outcome in CS_d produces a minimum satisfaction level of $k - d$ for every bargainer. Clearly,

$$CS_0 = \emptyset \subseteq CS_1 \subseteq CS_2 \subseteq \ldots \subseteq CS_k = K .$$

Next, define the (*bargaining*) *depth*, d^*, by

$$d^* = \min \{ d : CS_d \neq \emptyset \} .$$

Because of the chain of containment relations above, d^* is well-defined. The *Compromise Set*, $CS(A) = CS$, is the subset of K defined by $CS = CS_{d^*}$. Thus, the Compromise Set is the set of all alternatives that are among the first d^* in the ranking of every bargainer, where d^* is the smallest value that makes this set non-empty. In terms of satisfaction, every alternative in CS gives every bargainer a satisfaction level of at least $k - d^*$, and any alternative not in CS gives at least one bargainer a satisfaction level strictly less than $k - d^*$. This observation demonstrates that CS is identical to the Kant-Rawls Social Compromise Hurwicz and Sertel (1997).

Bounds on the size of $CS \subseteq K$ can be found easily.

Lemma A1. $1 \leq |CS| \leq \min\{d^*, n\}$.

Proof. Because CS is non-empty, $|CS| \geq 1$. Because every element of CS must appear among each bargainer's d^* highest ranked alternatives, $|CS| \leq d^*$. Because $CS_{d^*-1} = \varnothing$, every element of CS is the $d^{*\text{th}}$ entry in the ranking of some bargainer. There are only n bargainers, so $|CS| \leq n$ follows.

The bargaining depth, d^*, can be as low as 1, when all bargainers rank the same alternative first. The maximum value of d^* is given by

Theorem A1. $d^* \leq \lfloor k - k/n + 1 \rfloor$.

Proof. The first d entries of all n rows of A contain nd items, some of which may be duplicates. Each item is one of the k alternatives in K. By the pigeonhole principle, some alternative must appear at least $\lceil nd/k \rceil$ times.

Now suppose that $d > k - k/n$. Because $nd/k > n - 1$, some alternative must appear n times in the first d items of all rows. This implies that $d \geq d^*$. If $k - k/n$ is integral, we have shown that $d^* \leq k - k/n + 1$; if not, we have shown that $d^* \leq \lceil k - k/n \rceil$. The conclusion now follows easily.

We now demonstrate by example that the bound in Theorem A1 is tight. That is, for any n and k, it is possible to find an $n \times k$ matrix A such that the bargaining depth $d^* = \lfloor k - k/n + 1 \rfloor$. We take $K = \{1, 2, \ldots, k\}$.

If $n \geq k$, a matrix with the required property is easy to construct; simply choose A so that, for $i = 1, 2, \ldots, k$, $a_{ik} = i$. Thus, each of the first k rows of A has a different final entry, guaranteeing that $d^* = k$.

Assume that $n < k$ and let $p = \lceil k/n \rceil$. For $i = 1, 2, \ldots, n$, and $j = 1, 2, \ldots, k$, define the entries of A^{Ak} by

$$a^{Ak}_{ij} = [(i-1)p + j] \bmod k .$$

It is easy to verify that, for the matrix A^{Ak} so defined, $d^* = k - p + 1$. For example, if n = 3 and k = 11,

$$A^{A11} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 5 & 6 & 7 & 8 & 9 & 10 & 11 & 1 & 2 & 3 & 4 \\ 9 & 10 & 11 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{pmatrix},$$

so $d^* = 8$, whereas if $n = 3$ and $k = 12$,

$$A^{A12} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 1 & 2 & 3 & 4 \\ 9 & 10 & 11 & 12 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{pmatrix},$$

so $d^* = 9$.

Now we generalize the definition of fallback bargaining to q-approval fallback bargaining, where $1 \le q \le n$. As indicated in the text, the idea is to weaken the conditions for compromise so that acceptance by only q bargainers is required for a compromise. Therefore, define the set of *q-approval agreements of depth d* by

$$CS_d^q = \{ x \in K : |\{ i : j_i(x) \le d \}| \ge q \}$$

for $d = 0, 1, 2, \ldots, k$. Thus, CS_d^q is the set of all alternatives that are ranked among the top d by at least q bargainers. This definition generalizes fallback bargaining, which sets q = n. Thus, $CS_d = CS_d^n$. Clearly,

$$CS_0^q = \varnothing \subseteq CS_1^q \subseteq CS_2^q \subseteq \ldots \subseteq CS_k^q = K .$$

Define the (*q-approval fallback bargaining*) *depth*, d_q^*, by

$$d_q^* = \min \{ d : CS_d^q \ne \varnothing \}$$

Because of the chain of containment relations above, d_q^* is well-defined. The *q-approval Compromise Set*, CS^q is the subset of K defined by $CS^q = CS_d^q$ where $d = d_q^*$. Thus, the q-approval Compromise Set is the set of all alternatives that are among the d_q^* highest ranked for at least q bargainers, where d_q^* is the smallest value that makes this set non-empty.

We now characterize the q-approval Compromise Set in terms of satisfaction. For alternative $x \in K$, recall that bargainer i's satisfaction level is $h_i(x) = k - j_i(x)$. Define the q-satisfaction level of x by

$$h^q(x) = \max\left\{h : \left|\left\{ i : h_i(x) \geq h \right\}\right| \geq q\right\}.$$

Thus, for at least q bargainers the level of satisfaction at x is at least $h^q(x)$, but there are fewer than q bargainers whose level of satisfaction at x strictly exceeds $h^q(x)$. In other words, $h^q(x)$ is the minimum satisfaction level of the q most satisfied bargainers at x.

Theorem A2. Let $x \in K$. Then $x \in CS^q$ if and only if x maximizes $h^q(x)$. The maximum value of $h^q(x)$ is $k - d_q^*$.

Proof. For any $x \in K$, $x \in CS_d^q$ if and only if there are at least q different bargainers, i, for whom $j_i(x) \leq d$, i.e., $h_i(x) \geq k - d$. First suppose that $x \in CS^q$. Because $CS^q = CS_d^q$ for $d = d_q^*$, it follows that there are at least q different bargainers, i, for whom $h_i(x) \geq k - d_q^*$, so $h^q(x) \geq k - d_q^*$. Suppose that $h_q(x) = h > k - d_q^*$. Then there are at least q bargainers, i, for whom $h_i(x) \geq h$, i.e., $j_i(x) \leq k - h$, so $CS_{k-h}^q \neq \varnothing$. But $k - h < d_q^*$, and d_q^* is the minimum value of d for which $CS_d^q \neq \varnothing$. This contradiction shows that if $x \in CS^q$, then $h^q(x) = k - d_q^*$.

Now suppose that $x \notin CS^q$. Then there are fewer than q different bargainers, i, for whom $j_i(x) \leq d_q^*$, i.e., $h_i(x) \geq k - d_q^*$. Therefore, $h^q(x) < k - d_q^*$. We have shown that the maximum value of $h^q(x)$ is $k - d_q^*$, and this value is attained if and only if $x \in CS^q$.

Bounds on the size of $CS^q \subseteq K$ are easily found.

Lemma A2. $1 \leq |CS^q| \leq \min\{nd^*/q, n\}$.

Proof. The proof that $1 \leq |CS^q| \leq n$ is exactly as in Lemma A1. Every element of CS^q must appear at least q times among the d_q* highest-ranked alternatives of the n bargainers. Thus $q|CS^q| \leq nd_q$*, and the lemma follows easily.

The bargaining depth, d_q*, can be as low as 1, when there is one alternative that is ranked first by at least q bargainers. In fact, d_1* = 1 always. The maximum value of d^q* is given by

Theorem A3. $$d_q^* = \left\lfloor \frac{kq - k + n}{n} \right\rfloor.$$

Proof. As in Theorem A1, the pigeonhole principle shows that some alternative must appear at least $\lceil nd/k \rceil$ times in the first d entries of all n rows of A. If $d > k(q - 1)/n$, then $nd/k > q - 1$, which implies that $d \geq d_q$*. If $k(q - 1)/n$ is integral, this proves that $d_q^* \leq k(q - 1)/n + 1$; if not, if proves that $d^q* \leq \lceil k(q - 1)/n \rceil$. The conclusion now follows directly.

The construction given above, exemplified by A^{A11} and A^{A12}, can be used to demonstrate that the bound in Theorem A3 is the best possible.

We now show that, for any value of q, the *q-approval Compromise Set*, CS^q , contains only Pareto-optimal alternatives. Recall that alternative y is *Pareto-superior* to alternative x, written $y \succ x$, if and only if $j_i(y) < j_i(x)$ for all $i = 1, 2, \ldots , n$. If x has the property that no y exists such that $y \succ x$, then x is *Pareto-optimal*.

Theorem A4. If $x \in CS^q$, then x is Pareto-optimal.

Proof. We prove that if x is not Pareto-optimal, then $x \notin CS^q$. Assume that $y \succ x$, that the bargaining depth is $d_q^* = d$, and that $x \in CS_d^q = CS^q$.

Then there exists a set of bargainers, $B \subseteq \{1, 2, \ldots, n\}$, such that $|B| \geq q$ and $j_i(x) \leq d$ for all $i \in B$. But for every $i = 1, 2, \ldots, n$, $j_i(y) < j_i(x)$. Thus, for every $i \in B$, $j_i(y) \leq d - 1$. It follows that $CS_{d-1}^q \neq \varnothing$, contradicting the hypothesis that the bargaining depth is d. The conclusion now follows.

Now we provide a formal description of fallback bargaining with impasse. For each of the n bargainers, a preference ranking on $K \cup \{I\}$ is assumed, where I represents "Impasse" and K is the set of alternatives. (Note that $I \notin K$.) If $|K| = k$, then the bargainers' preferences are specified by an $n \times (k + 1)$ matrix $B = (b_{ij})$ such that, for $i = 1, 2, \ldots, n$, the i^{th} row of B is bargainer i's ranking (in descending order). For each bargainer, i, and alternative $x \in K$, denote by $j_i^B(x)$ the value of j satisfying $b_{ij} = x$, and by $j_i^B(I)$ the value of j satisfying $b_{ij} = I$.

Now, for $d = 0, 1, 2, \ldots, k$, define the set of *depth d agreements* as

$$CSI_d = \left\{x \in K : j_i(x) \leq \min\{d, j_i(I)\} \ \forall \ i = 1, 2, \ldots, n\right\}.$$

Note that CS_d is the set of all alternatives that are among the top d in the ranking of every bargainer, and that every bargainer prefers to I. Clearly,

$$CSI_0 = \varnothing \subseteq CSI_1 \subseteq CSI_2 \subseteq \ldots \subseteq CSI_k \subseteq K.$$

Note that $CSI_k = \varnothing$ is possible; it occurs when no alternative is preferred by every bargainer to I.

The *Compromise Set with Impasse*, *CSI*(*B*) = *CSI*, is defined by $CSI = \{I\}$ if $CSI_k = \varnothing$. Otherwise, define the *bargaining depth*, d^*, by

$$d^* = \min\{d : CSI_d \neq \varnothing\},$$

and define $CSI = CSI_{d^*}$. Thus, the Compromise Set with Impasse consists of either Impasse, or of all alternatives that are (i) preferred by every bargainer to I, and (ii) among the top d^* in the ranking of every bargainer; here, d^* is the smallest value that makes this set non-empty.

The next result connects fallback bargaining with and without impasse.

Theorem A5. In fallback bargaining with impasse, let the alternative set be K and let the $n \times (k + 1)$ matrix B represent preferences. Then $CSI(B) \neq \varnothing$, and, if $I \in CSI(B)$, then $CSI(B) = \{I\}$. If $x \in K$ and $x \in CSI(B)$, then $x \succ I$. Construct the $n \times k$ matrix A by deleting I from each row of B. If $x \succ I$ and $x \in CS(A)$, then $x \in CSI(B)$.

Proof. First note that, by construction, either $CSI(B) = \{I\}$ or $CSI(B) = CSI_d$ for some d such that $CSI_d \neq \varnothing$. Moreover, $CSI_d \subseteq K$. Thus $CSI(B) \neq \varnothing$, and either $CSI(B) = \{I\}$ or $CSI(B) \subseteq K$. Now suppose that $x \in K$ and $x \in CSI(B)$. Then x must precede I in the ranking of each bargainer, i, so $x \succ I$. Finally, suppose that $x \succ I$ and $x \in CS(A)$. If $x \notin CSI(B)$, then there exists $y \in K$ such that $y \succ I$ and $\max_i \{j_i^B(y)\} < \max_i \{j_i^B(x)\}$. But, for each i, $j_i^B(y) < j_i^B(I)$ and $j_i^B(x) < j_i^B(I)$, which implies that $j_i^A(y) = j_i^B(y)$ and $j_i^A(x) = j_i^B(x)$. It follows that $\max_i \{j_i^A(y)\} < \max_i \{j_i^A(x)\}$, contradicting the assumption that $x \in CS(A)$.

Corollary A1. All elements of $CSI(B)$ are Pareto-optimal.

Proof. First suppose that $CSI(B) = \{I\}$. By construction, there is no $x \in K$ such that $j_i^B(x) < j_i^B(I)$ for all $i = 1, 2, ..., n$, so I is Pareto-optimal. Otherwise, suppose that $x \in K$ and $x \in CSI(B)$. As demonstrated in the proof of Theorem A5, no alternative $y \in K$ Pareto-superior to x can exist. Thus x is Pareto-optimal.

Now we turn to the question of one bargainer's best response to another's ranking. The next theorem concerns the optimal ranking choice for bargainer $B1$ when $B1$ knows the ranking of its opponent, $B2$. For simplicity, we assume that $K = \{1, 2, ... , k\}$, and that $B1$'s true preference ranking over K is given (in descending order) by the permutation $E = \langle 1, 2, ... , k \rangle$. For now, we assume that $B1$ knows that $B2$'s ranking will be the permutation $b = <b_1, b_2, \ldots, b_k>$. $B1$ can submit any permutation $a = <a_1, a_2, \ldots, a_k>$.

Define a bargaining matrix $A(a, b)$ by constructing a $2 \times k$ matrix with first row a and second row b. Let $CS(a, b) = CS(A(a, b))$ denote the Compromise Set for this matrix. As a ranges over all permutations of K, many different Compromise Sets can arise. Denote this collection of subsets of K by $P(b) \subseteq 2^K$, the power set of K. $B1$'s objective is to pick a permutation a such that $CS(a, b)$ is most preferred within $P(b)$ (according to $B1$'s true preference ordering, E).

$B1$'s preference ordering, $E = <1, 2, \ldots, k>$ is defined on K rather than 2^K. To represent $B1$'s preferences on 2^K, and therefore on $P(b)$, we assume that

(A1) If $1 \leq r < s \leq k$, then $\{r\} \succ \{s\}$.

(A2) If $S \subseteq K$ and $|S| \geq 2$, then $\min\{r: r \in S\} \succ S \succ \max\{r: r \in S\}$.

where $S_1 \succ S_2$ means that $B1$ prefers S_1 to S_2. This preference ordering is a minimal extension from a complete order on K to a partial order on 2^K. Note that using this partial order, we cannot say whether $B1$ prefers $\{1, 3\}$ or $\{2\}$; we know only that $\{1\} \succ \{1, 2\} \succ \{2\} \succ \{2, 3\} \succ \{3\}$ and $\{1\} \succ \{1, 2\} \succ \{1, 3\} \succ \{2, 3\} \succ \{3\}$.

Given b, we say that a permutation $a = a^* = <a_1^*, a_2^*, \ldots, a_k^*>$ is a *best response* for $B1$ if $CS(a^*, b) = CS^*(b)$ is maximal within $P(b)$ according to this partial ordering. In general, it is possible for $P(b)$ to have many maximal elements; as will be seen below, however, the maximal subset (with respect to the partial order defined above) is always unique in this case. If so, we call the unique maximal subset $CS^*(b)$ an *optimal compromise* for $B1$, and call any a^* such that $CS(a^*, b) = CS^*(b)$ an optimal response for $B1$. If $CS(E, b) = CS^*(b)$, then $B1$ can do no better than to respond truthfully to b. We call b *incentive-compatible* (for $B1$) in this case.

For $t = 1, 2, \ldots, k$, define the *top-t set of b* by $U(b, t) = \{b_1, b_2, \ldots, b_t\}$. Thus $U(b, t)$ is the set consisting of the first t elements in $B2$'s permutation. Similarly, define the *top-t set of E* $= <1, 2, \ldots, k>$ by $U(E, t) = \{1, 2, \ldots, t\}$.

It is convenient to identify best responses according to the parity of the alternative set. The next theorem describes the case when the number of alternatives is odd.

Theorem A6. Fix b, and suppose that $k = 2h - 1$ where h is integral. Define $r = b_f = \min\{s: s \in U(b, h)\}$. Then $CS^* = \{r\}$, and b is incentive-compatible for $B1$ if either $r \geq f$ or $r < f$ and $U(b, f) \cap U(E, f) = \{b_f\}$. Otherwise, $V = \{s: r < s \leq f \text{ and } s \in U(b, f)\} \neq \emptyset$. In this case, b is not incentive-compatible for $B1$; to construct an optimal response, define $W = \{s \in K - U(b, f): s > f\}$. Then $|V| \leq |W|$, so given any enumerations of $V = \{v_1, v_2, \ldots\}$ and $W = \{w_1, w_2, \ldots\}$, a^* can be constructed as follows:

$$a_i^* = \begin{cases} w_s & \text{if } i = v_s \in V \\ v_s & \text{if } i = w_s \in W \text{ and } s \leq |V| \\ i & \text{otherwise} \end{cases}$$

Proof: First, note that the bound given by Theorem A1 is $d^* \leq h$. Thus the fallback process must end within h steps, and the best compromise set that $B1$ can hope to achieve is $\{r\} = \{b_f\}$. Note that $f \leq h$. We show that this Compromise Set can always be achieved by an appropriate choice of a^*. In fact, it is easy to verify that $CS(E, b) = \{r\}$ whenever $r \geq f$ or $r < f$ and $U(b, f) \cap U(1, f) = \{b_f\}$, so b is incentive-compatible for $B1$ in this case, and $CS^* = \{r\}$.

Now assume that $r < f$. It is easy to show that $U(b, f) \cap U(E, f) = V \cup \{b_f\}$. The case that remains to be settled occurs when $V \neq \emptyset$; if so, $CS(E, b)$ must contain an alternative other than r (it may contain r also).

We show how to achieve $\{r\} = CS^*(b)$. The overlap set V is the set of possible compromises that might supplant $\{r\}$; they follow r in $B1$'s ordering, but precede it in $B2$'s. The permutation a^* constructed as indicated is identical to $E = \langle 1, 2, \ldots, k \rangle$, except that all entries in V are interchanged with entries in W. This works because W contains only alternatives s that follow b_f in b, so cannot be in $CS(a^*, b)$. Also the entries of V and W cannot coincide, because $s \in V$ implies $s \leq f$, and $s \in W$ implies $s > f$.

It remains to show that $|V| \le |W|$. The elements of $U(b, f)$ that do not exceed f are r and the elements of V. Because $|V| + 1$ of the f entries of $U(b, f)$ are less than or equal to f, $f - |V| - 1$ entries must exceed f. But in all of K, only $2h - 1 - f$ elements exceed f. Thus W, which contains exactly the elements that exceed f and fall in K but not $U(b, f)$, must contain exactly

$|W| = 2h - 1 - f - [f - |V| - 1] = 2h - 2f + |V|$ entries. Because $f \le h$, if follows that $|V| \le |W|$, completing the proof.

When the number of alternatives is even, the situation is a little more complicated, as shown next:

Theorem A7. Fix b, and suppose that $k = 2h$, where h is integral. Define $r = b_f = \min \{s: s \in U(b, h)\}$. If $r < b_{h+1}$, then $CS^* = \{r\}$, and b is incentive-compatible for $B1$ if either $r \ge f$ or $r < f$ and $U(b, f) \cap U(E, f) = \{b_f\}$. Otherwise, b is not incentive-compatible for $B1$, and a^* as constructed in Theorem A5 produces $CS(a^*, b) = \{r\}$. If $r > b_{h+1}$, then $CS^* = \{r, b_{h+1}\}$ and b is incentive-compatible for $B1$ if and only if $r = h + 1$. Otherwise, an optimal response for $B1$ is

$$a_i^* = \begin{cases} b_{h+i} & \text{if } 1 \le i \le h \\ b_f & \text{if } i = h + 1 \\ b_{i-h-1} & \text{if } h + 2 \le i < h + 1 + f \\ b_{i-h} & \text{if } h + 1 + f \le i \le 2h \end{cases} .$$

Proof. Note that the bound given by Theorem A1 is $d^* \le h + 1$. The fallback process must end within h steps, producing a Compromise Set containing either one or two elements of $U(b, h)$, or end in exactly $h + 1$ steps, producing a Compromise Set containing b_{h+1} and one element of $U(b, h)$. It can be verified directly that, subject to these restrictions, the most preferred compromise set that $B1$ can achieve is $CS^* = \{r\}$ if $r = b_f < b_{h+1}$, and $CS^* = \{r, b_{h+1}\}$ if $r > b_{h+1}$. Again, the plan of the proof is to show that these possibilities can always be achieved.

First, if $r < b_{h+1}$, the proof of Theorem A6 can be mimicked to demonstrate that b is incentive-compatible for $B1$ if either $r \geq f$ or $r < f$ and $U(b, f) \cap U(E, f) = \{b_f\}$, and that otherwise b is not incentive-compatible for $B1$, and an optimal response a^* can be constructed exactly as in Theorem A6.

Now suppose that $r > b_{h+1}$. Then $B1$ prefers b_{h+1} to any of $b_1, b_2, \ldots, b_h$. It is easy to verify that if $B1$ submits the ordering $E = <1, 2, \ldots, k>$, then the result will be $CS^* = \{r, b_{h+1}\}$ if and only if $r = h + 1$, which means that the first h elements of b are the last h elements in $B1$'s true preference ordering.

If $r < h + 1$, it is easy to verify directly that $b_{h+1} \notin CS(E, b)$. Under the preference order on $P(b)$ defined above, $\{r, b_{h+1}\} \succ \{r\} \succeq CS(E, b)$. It is not difficult to verify directly that a^*, as constructed above, produces $CS(a^*, b) = \{r, b_{h+1}\}$.

The text contains examples showing the application of Theorems A6 and A7.

One example, given in the text for $k = 9$ (A[19]), is worth elaborating in general. It shows that the potential benefits of using a^* rather than the truthful ordering are unlimited. Suppose that $k = 2h - 1$, where h is integral, and that

$$b = <h - 1, h, \ldots, 2h - 3, 1, 2, \ldots, h - 2, 2h - 2, 2h - 1>$$

Note that $b_h = 1$, and that $CS(E, b) = \{h - 1\}$. By Theorem A6, $CS^*(b) = \{1\}$ because $r = 1$ and $f = h$. Using $V = \{h - 1, h\}$ and $W = \{2h - 2, 2h - 1\}$, the optimal Compromise Set $CS^*(b) = \{1\}$ can be achieved by the optimal response

$$a^* = <1, 2, \ldots, h - 2, 2h - 2, 2h - 1, h + 1, \ldots, 2h - 3, h - 1, h>$$

Note that the use of the best response improved the Compromise Set from $\{h - 1\}$ to $\{1\}$, relative to the truthful response, *E*.

We end this Appendix with a characterization of truthful equilibria, which are illustrated in the text. We continue with the assumptions introduced earlier to define a partial preference order for *B*1 on *P*(*b*), and apply them also to *B2* to produce a partial order on *P*(*a*). We use subscripts to indicate which partial order is being referred to; for instance, $s_1 \succ_a s_2$ means s_1 is preferred to s_2 according to ordering *a*—in other words, that s_1 precedes s_2 in *a*. Likewise, if $S \subseteq K$, then $\max_a\{S\}$ is the most-preferred element in *S* according the ranking *a*, and $\max_b\{S\}$ is the most-preferred element in *S* according the preference ranking *b*. As usual, the ranking is in descending order of preference, so alternatives earlier in the ranking are more preferred.

Let *CS**(*b*; *a*) represent the optimal Compromise Set for *B*1, based on preference permutation *a*, given that *B*2's ranking is *b*, and let *CS**(*a*; *b*) represent the optimal Compromise Set for *B*2, based on preference permutation *b*, given that *B*1's ranking is *a*. The pair of rankings (*a*, *b*) represents a Nash equilibrium if and only if $CS(a, b) = CS^*(b; a) = CS^*(a; b)$, in other words, if *a* is a best response to *b* for *B*1, given that *a* represents *B*1's true preference ranking, and *b* is a best response to *a* for *B*2, given that *b* represents *B*2's true preference ranking.

The next two theorems characterize all Nash equilibrium pairs of rankings. As usual, that the situation is more complicated when the number of alternatives is even.

Theorem A8. If $k = 2h - 1$ is odd, then (*a*, *b*) is a Nash equilibrium pair if and only if $\max_a\{U(b, h)\} = \max_b\{U(a, h)\} = c$ and, for some *d* satisfying $1 \leq d \leq h$, $U(a, d) \cap U(b, d) = \{c\}$.

Proof. From Theorem A6, $CS^*(b; a) = \max_a\{U(b,h)\}$ and $CS^*(a; b) = \max_b\{U(a, h)\}$, so the requirement that $\max_a\{U(b, h)\} = \max_b\{U(a, h)\}$ follows from the definition of Nash equilibrium. Assuming it is met, let the most-preferred common alternative be *c*.

If, for some *d*, $1 \leq d \leq h$, $a_d = b_d = c$, then Theorem A6 shows that neither bargainer can do better than to submit its true preference ordering. Otherwise, suppose that $b_d = c$ and $a_f = c$, where $f < d$. By

Theorem A6, $B2$ cannot do better than to submit b. To achieve c, it may be necessary for $B1$ to submit an ordering other than a; this occurs if and only if $V \neq \varnothing$, where

$V = \{s \in K: a_f \succ_a s \succeq_a a_d$ and $s \succ_b b_d\}$. But $c \in U(a, d) \cap U(b, d)$, and $U(a, d) \cap U(b, d) = \{c\}$ iff $V = \varnothing$. The situation is analogous if c appears earlier in b than in a.

Theorem A9. If $k = 2h$ is even, define $u_1 = \max_a \{U(b, h)\}$, $v_1 = \max_a \{U(b, h + 1)\}$, $u_2 = \max_b \{U(a, h)\}$, and $v_2 = \max_b \{U(a, h + 1)\}$. Then (a, b) is a Nash equilibrium pair iff either $u_1 = v_1 = u_2 = v_2 = c$ and, for some d, $1 \leq d \leq h$, $U(a, d) \cap U(b, d) = \{c\}$, or $u_1 \prec_a v_1$, $u_2 \prec_b v_2$, $u_1 = v_2$, $u_2 = v_1$, and $U(a, h + 1) \cap U(b, h + 1) = \{u_1, u_2\} = \{v_1, v_2\}$.

Proof. Based on Theorem A7, analogous to the relation of Theorem A8 to Theorem A6.

The text contains examples illustrating Theorems A8 and A9.