Stabilizing Power Sharing

Brams, Steven J. and Kilgour, D. Marc

New York University, Wilfrid Laurier University

August 2007

Stabilizing Power Sharing

Steven J. Brams
Department of Politics
New York University
New York, NY 10003
USA
steven.brams@nyu.edu

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

Prepared for delivery at the 2007 Annual Meeting of the American Political Science Association, August 30th – September 2nd, 2007.

Abstract

Power sharing is modeled as a duel over some prize. Each of two players may either share the prize in some ratio or fire at the other player—either in sequence or simultaneously—and eliminate it with a specified probability. If one player eliminates the other without being eliminated itself, it captures the entire prize, but the prize is damaged over time when there is shooting.

Simultaneous shooting, which is more damaging than sequential shooting, tends to induce the players to share the prize and expand their opportunities for sharing it. It was effectively implemented by the superpowers with the doctrine of "launch on warning" during the Cold War, and it was strengthened by the development of second-strike capability. Deterring terrorism has proved a different matter, because terrorists are difficult to detect and present few targets that can be damaged.

Stabilizing Power Sharing[1]

1. Introduction

Power sharing has been problematic from time immemorial. Children have difficulty sharing toys and desserts. Couples have difficulty dividing responsibilities.

In the corporate world, it is rare for two CEOs to share power without crossing swords. After a merger, quarrels between the CEOs of the merged companies are common; sometimes they become so fierce that one CEO is forced out. Such a power struggle is almost always detrimental to the new company, occasionally leading to its collapse.

At the national level, no country in the world officially has two presidents or two prime ministers. When two party leaders agree to share the prime ministership, then one typically holds this position for one period followed by the other's taking the reins for another period.[2]

When there is power sharing among political parties in parliamentary democracies because no party wins a majority of seats in the parliament, it is most often of cabinet ministries. Usually the largest party is awarded the prime ministership, and there is no simultaneous sharing of this prize.

At the international level, it is quite common for countries to rotate offices in an international organization. A new secretary-general of the United Nations never comes from the same country and almost never from the same region of the world as his or her

[1] We thank Eric S. Dickson for valuable comments on an earlier version of this paper.

[2] This happened, for example, when a national-unity government, comprising the two largest parties in Israel, assumed power over the 4-year period from 1983 to 1986. Itzhak Shamir of the Likud Party was prime minister for the first two years, and Shimon Peres of the Labor Party for the next two years.

predecessor, just as the presidency of the Council of Ministers of the European Union rotates every six months among its 27 members. Still, the largest countries in these organizations often exercise veto power—de facto or de jure—and sharing is anything but equal among the members of these organizations.

We focus in this paper on two-party power-sharing agreements and ask which factors make them stable. In a previous paper (Brams and Kilgour, 2007), we developed game-theoretic models in which players could agree to share power or engage in a duel. Each player had an unlimited number of bullets to expend, round by round.

By firing at an opponent and, with a specified probability, eliminating it, a player could capture all the assets. But because we assumed that the players were not perfect shots, shooting was not a surefire strategy to acquire these assets.

Ominously, we found that power sharing was almost never rational, however the assets were divided and however they were discounted in repeated play. Because the players almost always had an incentive to shoot, there was a "race to preempt."

The only way we found to slow down this race was to postulate that shooting would cause damage in each period that it occurred. But even this damage was often insufficient to deter the players from shooting, because they still received benefits in each period they survived.

If only one player survived, it benefited the most, because it received all the remaining assets. Because these assets were discounted or damaged more heavily the longer play continued, a player did best by eliminating its opponent early, which was abetted by its being a good shot.

In this paper, we assume that the game the duelists play is different from the ones we analyzed earlier. While repeated, it does not bestow payoffs on the players in each period that shooting occurs and neither is eliminated. Instead, there is a single prize, awarded at the end of play, which goes to

- *both players* if they agree to share it; or
- *one or neither player* if they refuse to share it and instead fire at each other until one or both is eliminated.

We consider two possibilities for shooting—that it may occur either sequentially or simultaneously. Although power sharing can occur for each possibility, the power-sharing region is considerably enlarged when shooting is simultaneous. Simultaneity also makes more sharing arrangements stable, so players have greater opportunity to design an agreement without fear that it will be abrogated.

2. Notation and Assumptions

Assume there are two players, P and Q. Power is a prize that both players may share at any time and has an initial value of 1.[3] If P and Q decide to share the prize, they do so in the ratio of $a : (1 - a)$, which is a ratio that we assume was set before play commenced. If the value of the prize when the players agree to share it is v, then P receives a payoff of av and Q a payoff of $(1 - a)v$.

Alternatively, P and Q may attempt to eliminate one another. If P fires at Q, Q is eliminated with probability p; if Q fires at P, P is eliminated with probability q. When a

[3] Power is often conceptualized as a relationship between players, not a good they may share. Because it is not apparent what sharing means in a power relationship, we posit a divisible good (prize) that the players agree to share or, by shooting, try to capture entirely.

player is eliminated, its payoff is 0. The survivor, if any, wins the entire prize. We assume that there is no disgrace or other penalty incurred from firing and missing an opponent.

Once started, firing proceeds in rounds as long as both players survive. If one or both players are eliminated, the game terminates, and the survivor, if any, receives the prize at that time.

In any round, both players have one opportunity to eliminate their opponent. A round of shooting in which neither player is eliminated reduces the value of the prize by a factor of $1 - s$, which reflects the damage caused by firing. Consequently, the prize is worth 1 in the first round, s in the second round, s^2 in the third round, and so on. If there are n rounds of fighting in which neither player is eliminated, and if the prize is then won during the $(n + 1)^{\text{st}}$ round, then it is worth s^n.

The payoff to a player is the expected value of the prize it receives. The players value nothing else, and firing has no cost.[4]

To avoid trivial cases, we usually assume that $0 < a < 1$, $0 < s < 1$, $0 < p < 1$, and $0 < q < 1$, and their values are common knowledge. While a may be related to the other parameters, including p, q, or s, we assume no specific relationship in our models. Instead, we identify the values of a (in terms of p, q, and s) that make sharing the prize—as opposed to fighting for it—a rational choice of the players.

Unlike our earlier models (Brams and Kilgour, 2007), we assume there are no interim rewards—in particular, there is no accumulation of payoffs, round by round, as

---

[4] This no-cost assumption differs from that in most economic models, in which players use up resources when they attack one another. We do not develop such a model here in order to focus on the conditions that discourage fighting when it is not costly. But cost considerations come into play indirectly—fighting makes the prize less valuable.

long as the players survive.[5] In particular, neither player receives anything until (i) each agrees to share the prize (once and for all), or (ii) at least one player is eliminated.

While time plays no direct role in our models, the players know that play cannot continue indefinitely (see note 4). The damage parameter, *s*, is effectively a discount parameter, whereby the prize shrinks in value as fighting continues. Consequently, even winning all of it in some later round will be less advantageous than sharing it at the start of play.

We turn next to assessing the effects of sequential versus simultaneous shooting. As we will show, simultaneous shooting is more likely to deter the players from firing, because it is more fearsome: It may cause more damage early; and it may eliminate both players on any round, which sequential shooting can never do.

3. Sequential Interaction

We assume the players act in sequence: Either the players agree at the outset to share the prize, or one of them fires at its opponent. If, say, P eliminates Q, P receives the prize, which has value 1. If P fails, Q responds by firing at P. If Q eliminates P, Q receives the prize, still worth 1. But if Q also fails, the players are in the same position as at the start, except that the value of the prize has been reduced from 1 to *s*.

We search for Nash equilibria in stationary strategies, which means that a player's strategy depends only on its strategic possibilities at the moment and not on the history of the players' interaction. Thus, a stationary strategy that calls for a player to try to

---

[5] If anything, costs rather than rewards accumulate as play continues. Firing uses up ammunition and other resources, which are not in reality unlimited. The models we develop probably apply best to situations in which P and Q have more or less equal resources, so a war of attrition does not favor either player. While fighting always ends in a finite number of rounds because *p* and *q* are positive, one cannot say exactly when it will end, except in probabilistic terms.

eliminate its opponent in the first round must, if both players survive the first round, call for the player to try to eliminate its opponent on the second round, and so on in future rounds.[6]

To determine whether sharing or firing is better for P, we calculate P's expected reward, V_P, if P fires at Q, noting that if Q survives, Q will fire back at P in the same round (Q has nothing to lose and possibly something to gain if it eliminates P). Because both players survive with probability $(1 – p)(1 – q)$, we have

$$V_P = p(1) + [(1 – p)q](0) + [(1 – p)(1 – q)](sV) = p + [(1 – p)(1 – q)](sV),$$

where the sV factor on the right side of the equation reflects the continuation of the game to a second round in which V is reduced to sV. If follows that

$$V_P = \frac{p}{1-[(1-p)(1-q)]s}. \tag{1}$$

P is rationally deterred from initiating the firing if and only if (iff) $V_P \leq a$, which is equivalent to

$$p \leq \frac{a-as(1-q)}{1-as(1-q)}. \tag{2}$$

[6] Why is this plausible? Because the only feature that has changed in the second and subsequent rounds is the value of the prize, which has decreased, so the strategic incentives remain the same because there is nothing in our model that relates the size of the prize to these incentives. To illustrate a nonstationary strategy, assume that after one round of firing, P chooses not to fire to try to induce its opponent to share the prize. Because P's behavior changes in the course of play, history matters, rendering its strategy nonstationary.

The fraction on the right side of (2) is the threshold value of p for deterrence to occur—that is, for P to prefer its share of the prize, a, to what it obtains, on average, from fighting.

Note that the numerator of the right side of (2) is $a[1 - s(1 - q)]$. Since

$$[1 - s(1 - q)] < [1 - as(1 - q)]$$

because $a < 1$, it follows that, independent of the values of s and q, if P is rationally deterred, then $p < a$. If $p \geq a$, (2) implies that a rational P will never be deterred from initiating the firing.

Similarly, Q will be rationally deterred iff its expected value, V_Q, is not greater than $1 - a$, the value it receives from sharing. Analogous to (2), the condition for deterrence of Q is

$$q \leq \frac{1 - a - s(1 - a)(1 - p)}{1 - s(1 - a)(1 - p)}. \tag{3}$$

Just as P is rationally deterred when the right side of (2) is less than a, Q is rationally deterred when the right side of (3) is less than $1 - a$. In particular, if $q \geq 1 - a$, a rational Q will never be deterred from initiating the firing.

Rewriting (3) as a condition on p and combining it with (2) shows that (2) and (3) both hold iff p satisfies

$$\frac{a - (1 - q) + s(1 - a)(1 - q)}{s(1 - a)(1 - q)} \leq p \leq \frac{a - as(1 - q)}{1 - as(1 - q)} \tag{4}$$

and, of course, $0 < p < 1$. The points (q, p) defined by these conditions are shown as the shaded region in Figure 1 for three cases: s approaches 0; $0 < s < 1$; and s approaches 1.

Figure 1 about here

Inequality (4) provides both lower and upper bounds on p. The upper bound on p always lies between 0 and a, and it is strictly decreasing in s and increasing in q. It approaches a as s approaches 0 or as q approaches 1, and it approaches 0 as s approaches 1 and q approaches 0.

The lower bound for p given by (4) is nonpositive when

$$q \leq \frac{(1-s)(1-a)}{1-s(1-a)},$$

which explains why the additional condition, $p > 0$, may come into play. When

$$\frac{(1-s)(1-a)}{1-s(1-a)} < q < 1-a,$$

this lower bound is positive, increasing in q and decreasing in s.

As q approaches 1, the numerator on the right side of (4) approaches a, and when $q = 1 - a$, the numerator on the left side of (4) equals 1. Thus, for example, as q approaches 0, P is rationally deterred from firing iff

$$0 < p \leq \frac{a - as}{1 - as}.$$

As Figure 1 shows, for any fixed (positive) value of s, deterrence is possible if p and q are sufficiently small. Deterrence is maximal when damage is nearly total (i.e., s is near 0), which occurs when $p < a$ and $q < 1 - a$. The rectangular area defined by these inequalities is greatest when $a = 1 - a = 1/2$, rendering the deterrence region a square. Thus, players that share the prize equally are most likely to be deterred from firing.

Deterrence is impossible if, when the players fire, no damage is inflicted because $s = 1$. For the players to be deterred from shooting, therefore, they must incur some damage from firing, and their probabilities of eliminating their opponents must not be too high.

In the special case when $a = 1/2$ and the players share the prize equally, the deterrence region—the set of (q, p) values where both players are rationally deterred from firing— is symmetric (it becomes a square as s approaches 0). The corner point of the deterrence region opposite the origin (0, 0) is (x, x), where

$$x = \frac{s - 1 + \sqrt{s - s^2}}{s}.$$

Note that x is a decreasing function of s, which approaches 0 as s approaches 1; it approaches 1/2 as s approaches 0. When $a = ½$ and s approaches 0, the area of the deterrence region is maximal at 1/4 of the (q, p) unit square. Thus, even in the best case of total damage and equal sharing, both players' shooting accuracies cannot exceed 1/2 for deterrence to occur.

4. Simultaneous Interaction

We now assume that the players act simultaneously (or that if one player fires first, its opponent can return fire, regardless of whether the first shot hits its mark). Thus, either the players agree at the outset to share the prize, or they fire at each other. In the latter case, it is possible for both shots to be successful, eliminating both players in any round, so each would receive a payoff of 0.

If the first shot is successful and a player is therefore eliminated, it would appear inconsistent to allow the eliminated player to return fire. However, there are instances of people who are fatally shot but, while taking their dying breath, manage to kill an assailant. At the international level, a “doomsday machine” also works in this manner, enabling state A to destroy B even as A itself is destroyed. By contrast, instantaneous reciprocation cannot happen in the sequential-interaction model, because an eliminated player cannot subsequently eliminate its opponent.

As in the sequential-interaction model, the value of the prize in the simultaneous-interaction model is reduced by the factor of $1 - s$ on each round if both players fire and neither is eliminated. Also as before, we restrict our analysis to stationary strategies.

To determine whether sharing or firing is better for P, we calculate P’s expected payoff, W_P, if P fires at Q. P will receive a positive payoff if P’s shot succeeds and Q’s (simultaneous) shot fails, whereas P will receive a payoff of 0 if Q’s shot succeeds. If neither player’s shot hits the mark, which will occur with probability $(1 - p)(1 - q)$, both players will survive and the game will continue to a new round:

$$W_P = [p(1 - q)](1) + q(0) + [(1 - p)(1 - q)]sW = p(1 - q) + [(1 - p)(1 - q)]sW.$$

This equation can be rewritten as

$$W_P = \frac{p(1-p)}{1-[(1-p)(1-q)]s}. \tag{5}$$

P is rationally deterred from initiating the firing iff $W_P \leq a$, which is equivalent to

$$p \leq \frac{a - as(1-q)}{(1-q) - as(1-q)}. \tag{6}$$

The fraction on the right side of (6) is the threshold value of p for deterrence. This threshold is always positive; it is less than 1 iff $a - as(1 - q) < (1 - q) - as(1 - q)$, which reduces to $q < 1 - a$. Hence, if $q \geq 1 - a$, P is rationally deterred from firing no matter what the value of p is.

Analogous to (6), Q is rationally deterred from firing iff

$$q \geq \frac{1 - a - s(1-a)(1-p)}{(1-p) - s(1-a)(1-p)}. \tag{7}$$

The threshold value of q, given by the right side of (7), is always positive, and it is less than 1 iff $p < a$. Hence, if $p \geq a$, Q is rationally deterred from firing no matter what the value of q is.

It is rational for P and Q to share the prize iff both (6) and (7) hold. Rewriting (7) as a lower bound on p (rather than an upper bound on q) shows that power sharing in the ratio $a : (1 - a)$ is rational for both players iff

$$\frac{a - (1-q) + s(1-a)(1-q)}{q + s(1-a)(1-q)} \leq p \leq \frac{a - as(1-q)}{(1-a) - as(1-q)}, \tag{8}$$

and, of course, $0 < p < 1$. The deterrence region, which are the points of the (q, p) unit square defined by (8), is shaded in Figure 2 for three cases: s approaches 0, $0 < s < 1$, and s approaches 1.

Figure 2 about here

Note that for any value of a, there are always some (q, p) values for which both players prefer to share the prize in the ratio $a : (1 - a)$ rather than fight. The deterrence region includes all points where $p \geq a$ and $q \geq 1 - a$; in particular, it includes points where the values of p and q are both near 1. Unlike the sequential-interaction model, both players benefit from sharing when they have high probabilities of eliminating each other.

The deterrence region also includes points where the values of q and p are near 0, but those points are much more confined, as Figure 2 makes clear. But as s falls, the damage caused by firing increases, and the deterrence region near $(q, p) = (0, 0)$ grows larger.

Figure 2 also shows that as s approaches 0, the deterrence region includes the rectangle with opposite corners $(0, 0)$ and $(1 - a, a)$, and the rectangle with opposite corners $(1 - a, a)$ and $(1, 1)$. In other words, in a broad band around the 45° line from (0, 0) to (1, 1), both players will be deterred.

Note that for any fixed s with $0 < s < 1$ (the middle case of both Figure 1 and Figure 2), the intersections of the curved lines in Figures 1 and 2 with the q- and p-axes are identical. This shows that at any (q, p) where deterrence is rational in the sequential-interaction model (Figure 1), it is also rational in the simultaneous-interaction model (Figure 2).

5. *How* Should Power Be Shared to Induce Stability?

We now take a different approach to power sharing, asking a design question: When power is to be shared in the ratio $a : (1 - a)$, what values of a render power sharing stable? More specifically, given p and q, what are the stabilizable values of a, if any, and for each of these, what values of s support power sharing? As we will see, the answers to these questions depend fundamentally on whether the interaction is sequential or simultaneous.

Sequential Interaction

Suppose that P and Q are interacting sequentially (SQ). Then P will rationally be deterred from initiating the firing iff $V_P \leq a$. From (1),

$$V_P = f_{SQ}(p,q,s) = \frac{p}{1-(1-p)(1-q)s} \leq a. \tag{7}$$

Analogously, Q will rationally be deterred from initiating the firing iff

$$V_Q = \frac{q}{1-(1-p)(1-q)s} \leq 1-a. \tag{8}$$

Inequality (8) is equivalent to

$$a \leq g_{SQ}(p,q,s) = \frac{1-q-(1-p)(1-q)s}{1-(1-p)(1-q)s}. \tag{9}$$

Combining (7) and (9), power sharing in the ratio a : (1 – a) is stable—neither P nor Q will initiate the firing—for all values of a that satisfy the double inequality,

$$f_{SQ}(p,q,s) \leq a \leq g_{SQ}(p,q,s). \tag{10}$$

Now suppose that $p > 0$ and $q > 0$ are fixed and consider the behavior of the functions, $f_{SQ}(p,q,s)$ and $g_{SQ}(p,q,s)$, as s increases from 0 to 1. It is easy to verify that

P's expected reward is bracketed by a lower bound of p and an upper bound that is a function of p and q,

$$f_{SQ}(p,q,0) = p \leq f_{SQ}(p,q,s) \leq \frac{p}{p+q-pq} = f_{SQ}(p,q,1), \tag{11}$$

for any value of s satisfying $0 \leq s \leq 1$. Furthermore, from (7), $f_{SQ}(p,q,s)$ is strictly increasing in s. From (8), $g_{SQ}(p,q,s)$ is strictly decreasing in s, and, analogous to (11),

$$g_{SQ}(p,q,0) = 1-q \geq g_{SQ}(p,q,s) \geq \frac{p-pq}{p+q-pq} = g_{SQ}(p,q,1) \tag{12}$$

for any value of s satisfying $0 \leq s \leq 1$.

Comparing the two right-hand expressions in (11) and (12), we find

$$g_{SQ}(p,q,1) = p-pq < p = f_{SQ}(p,q,1), \tag{13}$$

because of our assumptions that $p > 0$ and $q > 0$. Inequality (13) contradicts inequality (10), so (10) cannot be true when $s = 1$.

Thus, when there is no damage, there is no possibility of power sharing when interaction is sequential. One player will initiate the shooting, which will continue until one player is eliminated and the other player obtains all the (undamaged) value.

Note that the difference, $g_{SQ}(p,q,s) - f_{SQ}(p,q,s)$, is a strictly decreasing function of s, because both g_{SQ} and $-f_{SQ}$ are strictly decreasing functions of s. From (10) it follows that power sharing is possible if and only if this strictly decreasing difference is nonnegative, allowing for values of a that would stabilize power sharing. This implies that if power sharing is possible for some specific value of s, say $s = s_0$, then it is also possible for all $s < s_0$, and in particular for $s = 0$.

But from (11) and (12) we know that $g_{SQ}(p,q,0) - f_{SQ}(p,q,0) \geq 0$ iff $(1 - q) - p \geq 0$, or, equivalently, $p + q \leq 1$. Therefore, there is no possibility for power sharing (with sequential interaction) when $p + q > 1$. In other words, if the sum of the elimination probabilities is too high, each player will have an incentive to get in the first shot.

Next suppose that p and q satisfy $p + q = 1$. Then

$$g_{SQ}(p,q,0) - f_{SQ}(p,q,0) = 1 - q - p = 0,$$

which implies that power sharing is possible, but only for $s = 0$. In addition, because a must satisfy (10), power can be shared only in the ratio $a : (1 - a) = p : q$. In conclusion, power can be shared if $p + q = 1$, but only if damage is total ($s = 0$) and the power-sharing agreement exactly reflects the elimination-probability ratio ($p : q$).

The case $p + q < 1$ is all that remains. By (10), power-sharing can be stabilized for any value of s that satisfies $f_{SQ}(p, q, s) \leq g_{SQ}(p, q, s)$, which is equivalent to

$$s \leq \frac{1 - p - q}{(1 - p)(1 - q)} = s_{\max}(p,q).$$

If $s = 0$, power can be shared in the ratio $a : (1 - a)$ iff a satisfies the inequality $p \leq a \leq 1 - q$. But, as can be verified directly, if $s = s_{\max}(p, q)$, power must be shared in the ratio $a_0 : (1 - a_0)$, where

$$a_0 = a_0(p,q) = \frac{p}{p+q},$$

which is the limiting case discussed in the previous paragraph.

The possibilities for power sharing are illustrated in Figure 3. Note that all values of a such that $p \leq a \leq 1 - q$ induce stability if $s = 0$, but the

Figure 3 about here

interval of stabilizable values of a (shaded area in Figure 3) diminishes in length as s increases. When s reaches $s_{\max}(p, q)$, the interval contains only the single point $a_0(p, q)$, and it vanishes entirely as s increases further. In fact, it can be shown that the length of this interval decreases at an increasing rate as s increases.

We conclude that sequential interaction offers relatively few opportunities to stabilize power sharing. First, players will not be deterred from shooting unless their combined probabilities of eliminating their opponents on any round are relatively low; otherwise, each player will find it advantageous to try to eliminate its opponent at the start. Second, even when this condition is met, the ratio of their power shares, $a : (1 - a)$, must more or less reflect the ratio of their elimination probabilities, $p : q$, for the players to be deterred from firing; in fact, only this ratio stabilizes power sharing if $s = s_{\max}(p, q)$,

Finally, the damage caused by firing on any round must be substantial. Indeed, if the value of the prize that the players seek is relatively undiminished on each round they shoot (i.e., if s is high), power sharing may be impossible, even when all other conditions for stability are met.

Simultaneous Interaction

Now suppose that P and Q are interacting simultaneously. Then P will be rationally deterred from initiating the firing if $W_P \leq a$. From (5),

$$W_P = f_{SM}(p,q,s) = \frac{p(1-q)}{1-(1-p)(1-q)s} \leq a$$

and, analogously for Q,

$$W_Q = f_{SM}(p,q,s) = \frac{q(1-p)}{1-(1-p)(1-q)s} \leq 1-a.$$

The latter inequality is equivalent to

$$a \leq g_{SM}(p,q,s) = \frac{1-(1-p)q-(1-p)(1-q)s}{1-(1-p)(1-q)s}.$$

Therefore, power sharing in the ratio $a : (1 - a)$ is stable for all values of a that satisfy the double inequality,

$$f_{SM}(p,q,s) \leq a \leq g_{SM}(p,q,s). \tag{14}$$

Now suppose that $p > 0$ and $q > 0$ are fixed, and consider the behavior of the functions, $f_{SM}(p,q,s)$ and $g_{SM}(p,q,s)$, as s increases from 0 to 1. As in the case of sequential interaction, it is easy to verify that

$$f_{SM}(p,q,0) = p - pq \leq f_{SM}(p,q,s) \leq \frac{p-pq}{p+q-pq} = f_{SM}(p,q,1)$$

for any value of s satisfying $0 \le s \le 1$, and that $f_{SM}(p,q,s)$ is strictly increasing in s.

Similarly, $g_{SM}(p,q,s)$ is strictly decreasing in s, and

$$g_{SM}(p,q,0) = 1 - q + pq \le g_{SM}(p,q,s) \le \frac{p}{p+q-pq} = g_{SM}(p,q,1),$$

for any value of s satisfying $0 \le s \le 1$.

Observe that $g_{SM}(p,q,1) > f_{SM}(p,q,1)$, which implies that inequality (14) is true (for appropriate values of a) when $s = 1$. Moreover, $g_{SM}(p,q,s) - f_{SM}(p,q,s)$ is a strictly decreasing function of s. Therefore, for any values of p and q, power sharing (with simultaneous interaction) is possible for every value of s—that is, power sharing in some ratio is feasible, whatever the level of damage shooting causes.

As in the sequential-interaction case, the length of the interval of stabilizable values of a diminishes, at an increasing rate, as s increases. This is shown in Figure 4 for the same values of p and q that were used in Figure 3.

Figure 4 about here

The values of s and a that make sequential stabilization possible (darker shade in Figure 4) can be shown to be a subset of those that make simultaneous stabilization possible (lighter shade). Note in Figure 4 that

$$f_{SM}(p,q,1) = g_{SM}(p,q,1) \text{ and } f_{SQ}(p,q,1) = g_{SQ}(p,q,1).$$

However, the interval between these points stabilizes power sharing in the case of simultaneous interaction but not in the case of sequential interaction.

Clearly, simultaneous interaction is much more potent a tool than sequential interaction for stabilizing power sharing. More specifically,

- simultaneous stabilization is possible for any values of p and q, whereas sequential stabilization is possible only if $p + q \leq 1$;
- if $p + q \leq 1$, simultaneous stabilization is possible for every value of s, whereas sequential stabilization is possible only if $s \leq s_{\max}(p, q)$;
- if $p + q \leq 1$ and $s \leq s_{\max}(p, q)$, simultaneous stabilization produces a wider interval of values of a than does sequential stabilization.

The superior ability of simultaneous interaction to stabilize power sharing is made even more evident in Figure 5, which fixes $p = 1/2$ and asks how the stabilizable power sharing ratios depend on q. The figure includes three cases, $s = 0$ (total damage), $s = 1/2$, and $s = 1$ (no damage).

Figure 5 about here

Observe that as q increases from 0, the stabilizable values of a decrease. For example, when $s = 0$ and stabilization is simultaneous, values of a from 1/2 to 1 can be stabilized if $q = 0$, but at $q = 1$ the stabilizable values of a run from 0 to1/2. If interaction is sequential, the situation is bleaker: There are no stabilizable values of a when q exceeds 1/2. Thus, it is apparent that simultaneous interaction is far more efficacious at stabilizing power sharing than sequential interaction, especially when the elimination probability of a player increases.

Increasing s (i.e., decreasing damage) diminishes the possibility of stabilization in both the simultaneous and sequential cases. As suggested by the three cases in Figure 5, if a particular point (q, a) is stabilizable (for either sequential or simultaneous interaction) for any particular value of s, then it is also stabilizable for any smaller value of s. For instance, any point, (q, a), that is shaded (either light or dark) when $s = 1/2$ is also shaded when $s = 0$. Thus, the more the damage caused by shooting, the more players will try to avoid it.

6. Conclusions

Why are the incentives to share power in the simultaneous-interaction case greater than in the sequential-interaction case? The former allows for the possibility that both players will be eliminated and, consequently, receive none of the prize, whereas the latter model allows for at most one player to be eliminated. This makes the prospect of fighting more unsavory if interaction is simultaneous, raising the value to the players of sharing the prize.[7]

To deter players from firing and encourage power sharing, therefore, it helps if the players can respond rapidly, if not immediately, to firing by an opponent and so, potentially, wreak more damage. A hair trigger, despite the risks of accidental firing, therefore strengthens deterrence. So does the doctrine of "launch on warning," given

[7] The demise of dueling in the early 20th century seems to have been largely a function of the moral repugnance that came to be associated with it. But it also may have been due to the greater possibility that both players would be killed or wounded as pistols became more accurate. For a review of recent books on dueling, see Krystal (2007).

good intelligence and surveillance, because it enables the attacked player to retaliate before it is hit.

The near-simultaneity of possible retaliation by the superpowers during the Cold War arguably benefited deterrence, Wohlstetter's (1959) warning of the "delicate balance of terror" notwithstanding. Of course, mutual assured destruction (MAD) was never entirely assured, because the doomsday machines the superpowers put in place were not certain to work.

Failures of either command and control or political will were a constant concern, making the doomsday machines at best probabilistic (Brams, 1985, p. 36; Brams and Kilgour, 1988, pp. 50-52). However, as each side developed second-strike capability—primarily through its submarine-launched nuclear missiles, which could not be destroyed in a first strike despite the increased accuracy of ICBMs—MAD became more secure and, perhaps, less mad. Each side could ride out a first strike and still wreak destruction on the other side.

The simultaneous-interaction model mirrors this second-strike capability. Although firing may not literally be simultaneous, a player can respond to an attack, even if devastated by it, so a successful shot in simultaneous interaction does not "eliminate" an opponent entirely.

Put differently, even when great damage is inflicted on a player, it may be able to respond. What makes power sharing a rational strategy in this situation is the damage that both sides incur if both are eliminated at once (e.g., possibly a "nuclear winter" in the case of a nuclear exchange).

Unlike nuclear warfare, the damage caused by terrorist acts tends not to be highly destructive, except over long periods of time. Thus low damage, as well as sequentiality, may make terrorists reluctant to share power; instead, they do better by slowly wearing down the government. Indeed, the government may hasten its own demise if it fights back heavy-handedly, alienating the populace and, ultimately, losing its support if it is unable to detect and destroy many terrorist targets.

To increase the damage factor for terrorists, the best counterstrategy would seem to be to dry up their sources of support, especially financial, that derive from the populace. This, of course, is easier said than done. But we emphasize that the main lesson of our models is that the s factor—specifically, diminishing the value of the prize by making shooting (attacks) as damaging as possible—is the key to making power sharing attractive to both sides.

Can our models be extended to n-person power-sharing games, starting with truels, or 3-person extensions of duels (Kilgour and Brams, 1997; Bossert, Brams, and Kilgour, 2002)? The combinatorial possibilities of shooting rapidly multiply as the players increase, but so do the potential benefits of not shooting, so we think this question is well worth exploring in today's multipolar world.

References

Bossert, Walter, Steven J. Brams, and D. Marc Kilgour (2002). "Cooperative vs. Non-Cooperative Truels: Little Agreement, but Does That Matter?" *Games and Economic Behavior* 40, no. 2 (August): 185-202.

Brams, Steven J. (1985). *Superpower Games: Applying Game Theory to Superpower Conflict*. New Haven, CT: Yale University Press.

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

Brams, Steven J., and D. Marc Kilgour (1988). *Game Theory and National Security*. New York: Basil Blackwell.

Brams, Steven J., and D. Marc Kilgour (2007). "The Instability of Power Sharing." In Matthew Braham and Frank Steffen (eds.), *Power, Freedom, and Voting: Conceptual, Formal, and Applied Dimensions*. Heidelberg, Germany: Springer (forthcomng).

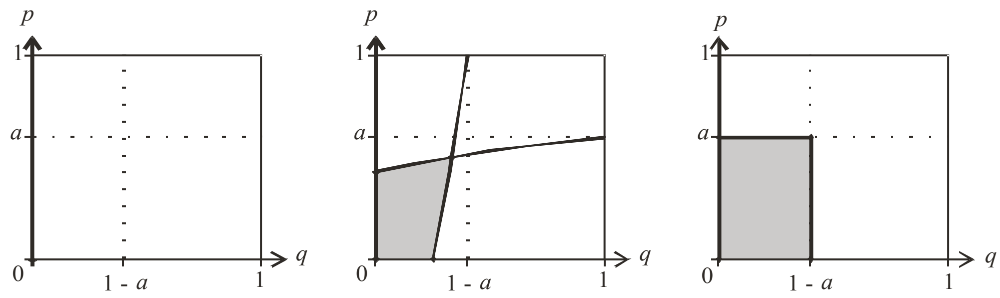
Kilgour, D. Marc, and Steven J. Brams (1997). "The Truel." *Mathematics Magazine* 70, no. 5 (December): 315-326.

Krystal, Arthur (2007). "En Garde!" *New Yorker*, March 12, pp. 80-84.

Wohlstetter, Albert (1959). "The Delicate Balance of Terror." *Foreign Affairs* 37 (January): 209-234.

Figure 1. Sequential-Interaction Model

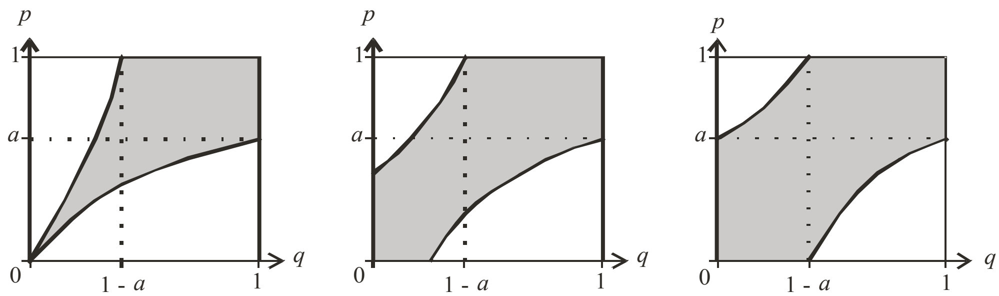
s = damage parameter
Power-sharing region shaded



s _ 1 (no damage) | $0 < s < 1$ | s _ 0 (total destruction)

Figure 2. Simultaneous-Interaction Model

s = damage parameter
Power-sharing region shaded



s _ 1 (no damage) | $0 < s < 1$ | s _ 0 (total destruction)

Figure 3. Sequential Stabilization (Shaded Area)

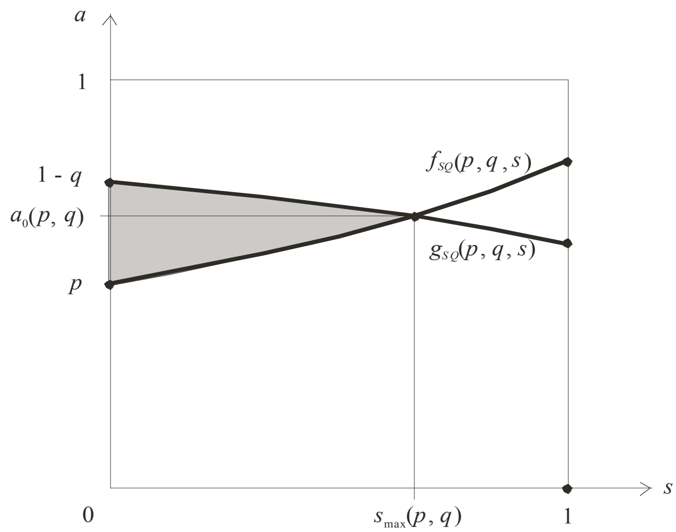


Figure 4. Comparison of Simultaneous Stabilization (Light Shading) with Sequential Stabilization (Dark Shading)

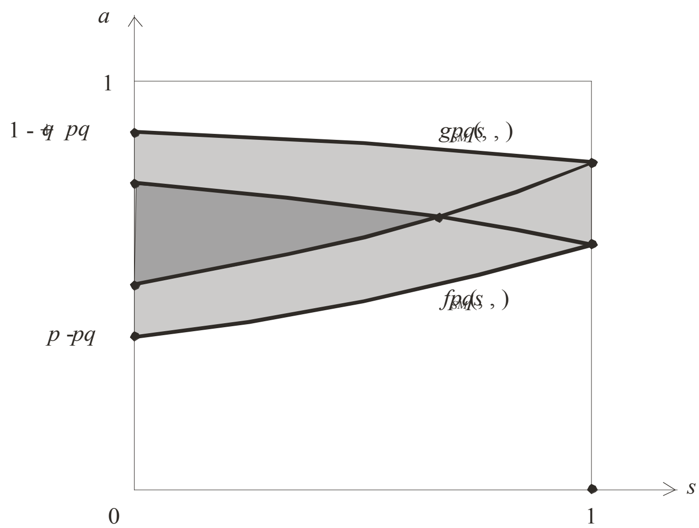
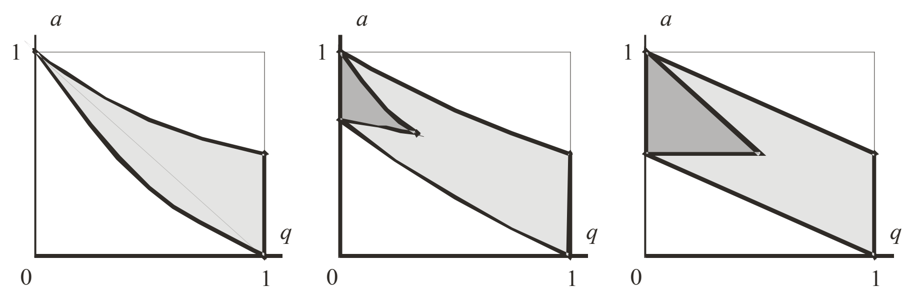


Figure 5. Stabilizable Values of a when $p = 1/2$

Simultaneous Stabilization: Light Shading
Sequential Stabilization: Dark Shading



s ?1 (no damage) | $0 < s < 1$ | s ?0 (total destruction)