Wilfrid Laurier University

Scholars Commons @ Laurier

Mathematics Faculty Publications   Mathematics

Spring 2001

The Impact of Conventional Force Reductions on Strategic Deterrence: A Game-Theoretic Analysis

D. Marc Kilgour
Wilfrid Laurier University, mkilgour@wlu.ca

Frank C. Zagare
University at Buffalo, the State University of New York

Follow this and additional works at: https://scholars.wlu.ca/math_faculty

Recommended Citation

Kilgour, D. Marc and Zagare, Frank C., "The Impact of Conventional Force Reductions on Strategic Deterrence: A Game-Theoretic Analysis" (2001). *Mathematics Faculty Publications*. 1.
https://scholars.wlu.ca/math_faculty/1

The Impact of Conventional Force Reductions on Strategic Deterrence: A Game-Theoretic Analysis

D. MARC KILGOUR

Wilfrid Laurier University

and

FRANK C. ZAGARE

University at Buffalo, The State University of New York

[**Abstract:** Many game-theoretic analyses of deterrence confirm the commonsense view that what determines whether a defender can effectively deter a challenger from an unwanted action is (1) the challenger's perception of the level of punishment that the defender will be able to impose on the challenger should it take the action, and (2) the challenger's level of belief about the likelihood of the defender actually carrying out this punishment. Reduction of the defender's forces may affect both the defender's ability to retaliate and its perceived willingness to do so. Game-theoretic methods are used to assess how the limits on both of these parameters are related, subject to the condition that deterrence remains effective. The results indicate that the defending side can often make do with smaller forces, provided its (apparent) resolve is high. But force structure is important—the models suggest that implementation of an "all-or-nothing" deployment (as called for by a doctrine of massive retaliation, for example) may reduce not only costs, but also deterrence effectiveness.]

Introduction

Since the end of the Cold War, the breakup of the Soviet Union, the disintegration of the Warsaw Pact, and the expansion of NATO, downsizing conventional military forces is all the rage. The reasons are primarily economic, though some would argue that in the United States military cutbacks have served political objectives as well as producing a "peace dividend." Force changes since the late 1980s have been substantial, as both Russia and the United States reduced military spending, cut back on ground forces, and moved to a greater dependence on "all-or-nothing" strategic-level threats to ensure security (Gacek, 1994). The current sweeping review of American defense policy by the Bush administration notwithstanding, there is no good

59

reason to believe that this trend will not continue—in the United States, in Russia, and elsewhere.

Yet, at the same time, these same powers are reducing their nuclear forces, both unilaterally and as a consequence of treaties like START and INF. A 1997 report of the National Academy of Sciences urged additional dramatic cuts in the two countries' nuclear arsenals, suggesting that they should eventually stabilize around 300 warheads each. The Academy also called on the United States to make a no-first use pledge, restricting the role of nuclear weapons to deterring or responding to nuclear attacks: "This country should no longer threaten to respond with nuclear weapons against attacks by conventional, chemical, or biological weapons."

Are these trends contradictory? Can downsized tactical forces maintain national security? The benefits of reductions in the level of resources consumed by a state's military can be assessed only in comparison to the costs of these reductions in terms of the state's goals, interests, and assets. In our view, military forces have two major (non-domestic) purposes—to fight, and to deter other states from unwanted acts by threatening to fight in response. So to judge whether cuts to a military are a good idea, it is essential to understand how much (if at all) those cuts reduce the military's ability to fulfill these purposes.

It is not our intention to address here questions of the fighting ability of reduced forces. This is not to say that fighting ability is unimportant, or easy to estimate. Clearly the capacity of a military depends in large measure on its level of resources; nonetheless, it may sometimes be possible to downsize forces while maintaining their fighting effectiveness. Savings in personnel costs, for example, often follow from technical developments, as when bombers were replaced by missiles, observers by radar, or battleships by guided-missile cruisers.

We pose a different question: Even if the reduced force would actually be as effective when engaged in combat, would its effectiveness as a threat be compromised by downsizing? In other words, does the act of reducing military forces make them less effective as a deterrent?

Our question, we believe, is a practical one. Downsizing the military is rarely an issue if

war is ongoing or imminent. When the actual fighting ability of the military is paramount, then its resource base is usually maintained, and cutbacks in any component are usually more than balanced by growth elsewhere. But overall cutbacks can become an issue whenever the military is not actually engaged in combat—no matter how vital the need to guard state security by threatening to respond to incursions by other states.

There are two ways that downsizing forces could diminish threat effectiveness. First, deterrence works because the threatenee fears that costs would be inflicted on it should hostilities break out. Knowledge that the threatener is reducing its forces may change the threatenee's estimate of the level of costs it would suffer should it provoke conflict. Note that it is the threatenee's *perception* of costs that is important here.

Second, even if ability to inflict costs is viewed as unaltered, the deliberate shrinking of the military resource base may be perceived as a signal of reduced willingness to commit forces to battle. Hitler, for example, seems to have drawn this inference from Britain's unilateral disarmament during the 1930s. In other words, downsizing one's forces may make the threat to use those forces less credible to a potential adversary. Thus the adversary's calculus may be altered both by the reduction of the perceived likelihood of conflict following the proscribed action, and the reduction of the adversary's estimate of the damage it would suffer should the conflict actually occur.

Our objective here is to explore the relationship among apparent willingness to fight, perceived ability to damage an adversary in a fight, and deterrence effectiveness, and to study whether and how this relationship changes in the context of downsizing. Our investigation is relevant in light of the apparent contradiction between the recommendations of the National Academy of Sciences and the opinion, held by many classical deterrence theorist, that overkill capability is necessary for deterrence success. (See, for instance, Intriligator and Brito, 1984.) We first discuss briefly the meaning of deterrence, and exhibit a particularly simple condition for

deterrence effectiveness that is robust across many models.[1] Our deterrence effectiveness condition is quite explicit, and provides us with insight into the details of the tradeoff among perceived capabilities, perceived willingness to retaliate, and deterrence effectiveness. We use this condition to draw conclusions about the risks of downsizing forces, and the parameters that determine whether the downsizing path is safe for the threatener.

Next, we introduce the *Rudimentary Asymmetric Deterrence Game*, a very simple model of the interaction of two states in which deterrence may or may not succeed. We demonstrate that the deterrence effectiveness condition applies in this model. We also review more complex models to which the deterrence effectiveness condition also applies. Then, to explore further how the values of the threatener affect deterrence success, we adjoin two models of these values to the Rudimentary Asymmetric Deterrence Game. This step allows us to model the relationship of changes in the threatener's values caused by downsizing with the credibility requirements for deterrence effectiveness.

Our study leads us to support the proposal of the National Academy of Sciences and other policies consistent with a posture of minimal deterrence, defined as "the retention of only enough nuclear weapons to provide an assured destruction capability" (Kegley and Wittkopf 1989: 351). We argue that such a policy, properly implemented, would maintain sufficiently credible retaliatory threats, and impose sufficiently high costs on an aggressor, to stabilize relationships and ensure the success of deterrence.

We also offer comment on a deployment policy that is often associated with downsizing. In general, the policy of "massive retaliation" is to threaten a disproportionately great response to any provocation. This policy has been justified by the argument that any reduction in threat credibility will be more than compensated by the risk of extreme damage—from large-scale

---

[1] The literature on deterrence is voluminous. Representative of the policy debates are Curtis (2000); Gholz and Sapolsky (1999/2000); Paul, Harknett and Wiltz (1998); Quinlan (2000/2001); Sanger and Eckholm (1999); Tammen et al. (2000); and Turner (1999). Recent empirical studies include Danilovic (2001a, 2001b) and Harvey (1998). For a literature review, see Huth (1999) or Zagare (1996).

response by strategic nuclear forces, for example. Moreover, it may sometimes be possible to restructure military forces to provide the capability of inflicting much greater damage using reduced forces, connecting massive retaliation with downsizing. We studied massive retaliation earlier (Zagare and Kilgour, 1993b), and showed that certain strategic considerations imply that it is usually ineffective. We discuss here the relevance of our conclusions to the question of how to downsize forces.

What Makes Deterrence Effective?

In our view, a very simple framework must support any deterrence model. At minimum there must be two sides, here called *Challenger* (Ch) and *Defender* (Def), and (at least) three possible outcomes, *Status Quo* (SQ), *Challenger Wins* (CW), and *Conflict* (Conf). At the *Status Quo*, Defender holds some object of value or privilege, the prize (or pawn), that Challenger wishes to obtain for itself. Challenger can either accept the *Status Quo*, or initiate a confrontation with Defender. If Challenger accepts it, then *Status Quo* is the final outcome. But if Challenger initiates, Defender can either concede the prize to Challenger, producing *Challenger Wins*, or resist the transfer of the prize, producing *Conflict*.

We assume that, in order to retain the prize, Defender has adopted a "deterrence policy." That is, Defender has threatened to respond to any initiation by Challenger, so that the consequence of initiation would not be *Challenger Wins*, but *Conflict*. We use the term "credibility" to refer to the likelihood that this threat would in fact be carried out.[2] Since credibility is associated with a threat that Defender is assumed to have made, we refer to it as "Defender's credibility."

There are several further conditions that this modeling framework must satisfy in order to

---

[2] This conception of deterrence applies whether Defender's threatened response is automatic or deliberate. All of our models assume that response is a choice completely controlled by Defender who, like Challenger, is a rational decision-maker, making choices—including carrying out its threat—in what it sees as its own best interests. For models in which Defender's response is automatic, but there is some doubt as to whether it will actually occur, see Brams and Kilgour (1988).

represent deterrence situations meaningfully. Challenger must prefer *Challenger Wins* to *Status Quo* to *Conflict*, for only in this case would Challenger choose to initiate if credibility is low enough, but not if credibility is high enough. To be more specific, initiation would make Challenger better off if credibility is low, because low credibility means that resistance is unlikely; thus, initiation would likely bring Challenger its most preferred outcome, *Challenger Wins*, rather than the middling outcome, *Status Quo*. If Challenger initiates, we say that Defender's deterrence policy has failed, or was ineffective. But when credibility is high, Challenger would be better off to remain at *Status Quo*, because to initiate would entail a high risk of ending up at its least preferred outcome, *Conflict*. If Challenger does not initiate, we say that Defender's deterrence policy has succeeded, or was effective.

Many models consistent with this framework, including some that are very simple, produce similar conclusions about whether deterrence will succeed. Across a wide range of models, a policy of deterrence succeeds if and only if the following "deterrence effectiveness condition" holds true:

$$\text{Def's credibility} \geq \frac{\text{Ch's gain at CW}}{\text{Ch's gain at CW + Ch's cost of Conflict}} \qquad (*)$$

Of course, the actual model being analyzed determines the precise meanings of the quantities in (*), but these meanings have much in common, and the commonalities can be understood in the context of the modeling framework.

For instance, the gain and loss appearing on the right side of (*) are to be measured relative to the *Status Quo*. (The fundamental requirements that Challenger prefer *Challenger Wins* to *Status Quo*, and *Status Quo* to *Conflict*, guarantee that both the gain and the loss are positive.) To see that (*) is plausible, note, for instance, that if Challenger has a great deal to lose at *Conflict*, then "Challenger's cost of Conflict" is large, so the denominator of the fraction

is large, making the fraction itself small. It follows that when Challenger fears *Conflict*,

deterrence can be effective even when Defender's credibility is low. On the other hand, if

possession of the disputed prize is worth a great deal to Challenger, then "Challenger's gain at

Challenger Wins" is large, increasing the value of the fraction and implying that deterrence is

effective only for high values of Defender's credibility.

The deterrence effectiveness condition, (*), thus captures what we think of as the

"commonsense" view of deterrence—a policy of deterrence is effective if and only if the gain

that results from initiation (providing there is no response) is small enough, or the loss that might

be triggered by initiation (when there is a response) is great enough. All of the models to which

(*) applies assume that the players do not know for certain whether there will be a response

(Zagare and Kilgour, 2000). (There are other uncertainties in some models, such as whether

Challenger will counter-respond, or which response option Defender will choose.) Note that (*)

is really a comparison of Challenger's beliefs about the likelihood of a response to initiation (left

side) and of Challenger's assessments of the potential gains and losses that might arise from

contesting the *Status Quo* (right side).

Modeling the Effects of Force Reductions

In the introduction, we expressed our view that force reductions by a Defender affect the

logic of deterrence in two ways, (1) by reducing Defender's credibility, and (2) by decreasing

Challenger's estimate of the cost of conflict. We now use our general condition for deterrence

effectiveness, (*), to assess these effects more systematically.

Figure 1 applies (*) to identify when deterrence is effective, as a function of Challenger's

cost of conflict and Defender's credibility. Notice that credibility lies between 0 and 1.[3] If the

cost of conflict to Challenger is near zero, then only a Defender with very high credibility can

---

[3] We will later identify Defender's credibility with the *a priori* probability that Defender prefers *Conflict* to *Challenger Wins*. A probability must be a number between 0 and 1.

deter. But if the cost of conflict to Challenger is very high, then even a low level of credibility is sufficient to deter.
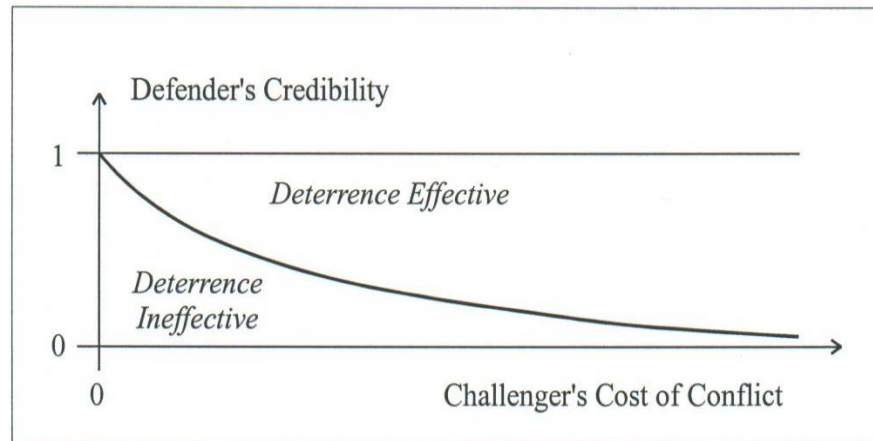

Figure 1: Cost of Conflict, Credibility, and Deterrence Effectiveness: The Trade-Off

Figure 1 plainly illustrates the rationale supporting a policy of minimum deterrence. A particular deterrence situation has a Defender credibility value and a Challenger cost of conflict that together determine a point in the region shown. If the point lies above the curved line, deterrence works. If not, it fails. And when deterrence fails, it fails completely. Thus, if deterrence is already effective, then Defender cannot possibly improve its position by making its response more credible, or raising the costs of conflict to Challenger. Once safely past the minimum, further expenditures by Defender are unproductive.

As discussed above, military cutbacks by Defender can have two relevant effects: they reduce (Challenger's estimate of) Challenger's cost of conflict, and they may reduce Defender's credibility. If a deterrence policy is effective, then the situation is described by a point located in the upper ("Deterrence Effective") zone. When Defender downsizes, the point describing the situation moves simultaneously to the left and downward. Clearly, there is a risk of crossing into the lower ("Deterrence Ineffective") zone. Note that if the line is crossed, then deterrence will

66

surely fail, but if not, then there is no risk of deterrence breakdown. Thus, we predict threshold effects—arms reductions will have negligible effects until a line is crossed, or is perceived to have been crossed, when the policy of deterrence will suddenly become ineffective.

Recall that, with suitable interpretations, (*) applies across a wide variety of models of deterrence. Our conclusions are based on Figure 1, which in turn is based on (*), and, therefore, are applicable to all of the deterrence models in this range. In particular, our prediction of threshold effects in downsizing applies in a wide range of deterrence situations including direct and extended deterrence relationships, and mutual and unilateral deterrence situations.

A Very Simple Deterrence Model

To gain more insight into the interpretation and significance of the deterrence effectiveness condition (*), we must give specific meaning to its terms. We will do this in the context of a very simple model of a deterrence situation. After showing that (*) applies to this model, we will compare it to other more complex models to which (*) also applies. Later, we will express our very simple model as an incomplete information game and introduce two information models that apply to it, providing us with two different viewpoints on deterrence effectiveness.

The Rudimentary Asymmetric Deterrence Game[4] is the simple extensive game shown in Figure 2. Its two players, Defender (Def) and Challenger (Ch), are assumed to have von Neumann-Morgenstern utilities for the game's three possible outcomes, which are denoted SQ (*Status Quo*), CW (*Challenger Wins*), and Conf (*Conflict*). Defender's utility at outcome K is denoted by d_K, and Challenger's by c_K.

---

[4] For a discussion and analysis of this game, see Zagare and Kilgour, 2000: Chapter 3 and Appendix 3.

Challenge

Initiate   Not
                 Initiate

Status Quo   **Defender**
(SQ)

         Not   Resist
         Resist

Challenger   *Conflict*
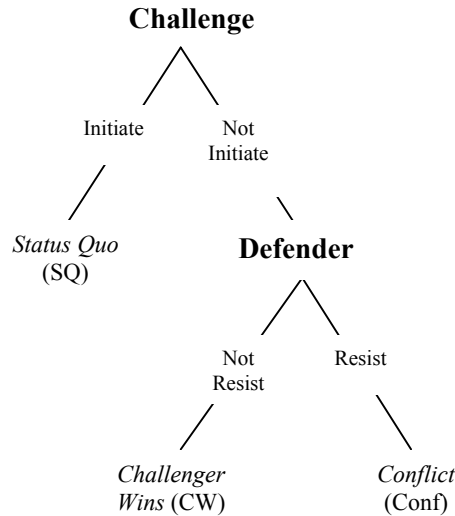*Wins* (CW)   (Conf)

Figure 2: Rudimentary Asymmetric Deterrence Game

Our requirement above on Challenger's preference ordering over the three possible outcomes corresponds to the condition that Challenger's utilities satisfy

$$c_{CW} > c_{SQ} > c_{Conf}$$

Thus, in this model, "Ch's gain at CW" is simply $c_{CW} - c_{SQ}$, and "Ch's cost of Conflict" is $c_{SQ} - c_{Conf}$. The right-hand side of inequality (*) then simplifies to

$$\frac{c_{CW} - c_{SQ}}{(c_{CW} - c_{SQ}) + (c_{SQ} - c_{Conf})} = \frac{c_{CW} - c_{SQ}}{c_{CW} - c_{Conf}}$$

We now show that the deterrence effectiveness condition (*) applies to the Rudimentary Asymmetric Deterrence Game. To do so, we identify Defender's credibility, the left side of (*), with the probability that Defender chooses Resist, rather than Not Resist, at its decision node. (See Figure 2.) If Challenger chooses Not Initiate, then the outcome of the game is *Status Quo*, and Challenger's utility is c_{SQ}. If Challenger chooses Initiate, then the outcome of the game, and

68

therefore Challenger's utility, depends on whether Defender chooses Resist or Not Resist. For convenience, denote Defender's credibility by p_H. Then Defender chooses Resist with probability pH, and Not Resist with probability $1 - p_H$. It follows that Challenger's expected utility if it chooses Initiate is

$$p_H \, c_{Conf} \; + \; (1 - p_H) \, c_{CW} \, .$$

Thus, Challenger is at least as well off to choose Not Initiate as Initiate, and the deterrence policy is effective, if and only if

$$c_{SQ} \; \geq \; p_H \; c_{Conf} \; + \; (1 - p_H) \, c_{CW} \, .$$

Manipulation of this inequality, making use of the fact that $c_{CW} > c_{SQ} > c_{Conf}$, produces

$$p_H \; \geq \; \frac{c_{CW} - c_{SQ}}{c_{CW} - c_{Conf}}$$

which, as already noted, is equivalent to (*) for the Rudimentary Asymmetric Deterrence Game. As an aside, we mention here some other game models in which (*) is also necessary for deterrence. The deterrence effectiveness condition applies in Kilgour and Zagare (1991, 1994b) and Zagare and Kilgour (1993a, 1993b, 1995, 1998, 2000), all of which are related to, but more complex than, the Rudimentary Asymmetric Deterrence Game. To suggest the flavor of (*) in these more complex models, we discuss one of them briefly. The *Asymmetric Escalation Game* (Zagare and Kilgour, 1994b, 1998) allows Defender to respond at either the tactical level or the strategic level; if the former, Challenger can counter-respond to the strategic level. A necessary condition for deterrence effectiveness in that game is

$$p_{Tac} \geq \frac{c_{DC} - c_{SQ}}{c_{DC} - c_{EE+}} \, ,$$

where the "credibility" p_{Tac} equals the probability that Defender prefers a tactical-level conflict to giving in to a Challenger who has initiated a sub-strategic conflict. This outcome, called DC, corresponds to *Challenger Wins*. The outcome SQ corresponds to *Status Quo*, and $c_{\text{EE+}}$ is Challenger's value for an all-out (strategic-level) conflict (that it would prefer to capitulation when Defender has escalated the conflict to the strategic level).

Thus, the deterrence effectiveness condition (*) applies across a wide variety of conceptions of deterrence. This robustness justifies its use in our study of the effects of downsizing.

Now, to explore further the consequences of downsizing, we now complete the formulation of the Rudimentary Asymmetric Deterrence Game as a game of incomplete information, and exhibit its solutions. Because this game is such a simple model, it is a very useful vehicle for interpreting the effects of force reductions.

To specify the incomplete information in the Rudimentary Asymmetric Deterrence Game as a game of incomplete information is to account for the choices made by Defender. We first model Defender's utilities, allowing us to associate Defender's credibility with Defender's preference for intervention (Zagare and Kilgour, 2000). We then identify all equilibria (technically, perfect Bayesian equilibria) of the game.

First, we always assume that Defender prefers to retain the prize. This means that Defender prefers *Status Quo* (SQ) to *Challenger Wins* (CW), or, in utilities,

$$d_{\text{SQ}} > d_{\text{CW}} .$$

Next, we assume that Defender prefers *Status Quo* to *Conflict* (Conf), for a Defender who prefers otherwise would want to induce conflict.

For reasons that will become clear below, we write

$$d_{SQ} > d_{Conf} .$$

We cannot say, however, whether Defender prefers *Conflict* or *Challenger Wins*; we must allow for both possibilities.

A detailed model accounting for Defender's value for *Conflict* would include many considerations, such as Defender's view of the likelihood that it will prevail, the costs of fighting, the possible damage it will suffer as a result of fighting, and the value of the prize (Kilgour and Zagare, 1994a). Because at least some of these quantities are generally unknown to Challenger, Defender's opponent, we now model the Rudimentary Asymmetric Deterrence Game as a game of incomplete information.

Thus, we represent Defender's value for *Conflict* as a random variable D_{Conf} drawn from a probability distribution as shown in Figure 3. Of course, our model is a model of one-sided incomplete information; so after a random value for D_{Conf} has been selected, it is reported to Defender but not to Challenger. Thus, both players know the distribution from which D_{Conf} is drawn, but only Defender knows its actual value.
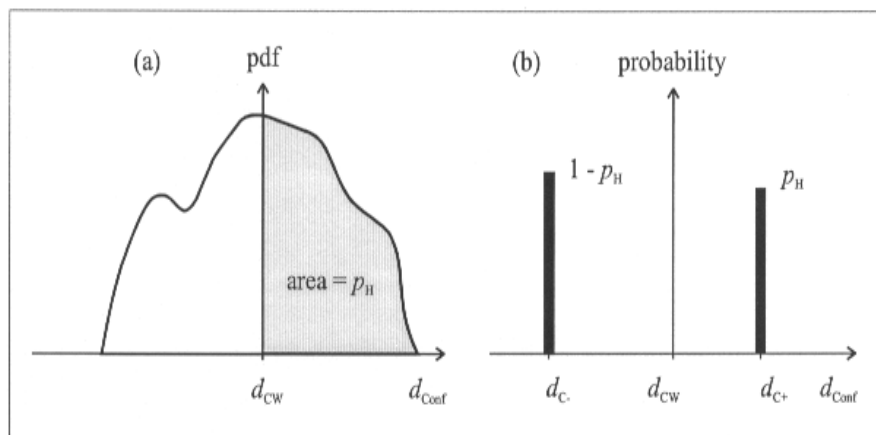

Figure 3: Model's of Defender's value for *Conflict* (D_{Conf}) in Rudimentary Asymmetric Deterrence Game

71

Figure 3a represents a general model in which d_{Conf} is a continuous random variable that takes all of its values, with certainty, below d_{SQ}. In fact, what Figure 3a shows is the probability density function of D_{Conf}.[5] A much simpler model of Defender's value for the *Conflict* outcome is shown in Figure 3b. Here, D_{Conf} is treated as a binary random variable.[6] Figure 3b shows a probability function representing the two possible values of D_{Conf}, and their probabilities, as follows:

$$D_{\text{Conf}} \quad = \quad \begin{cases} d_+ \ \text{ with probability } \ p_{\text{H}} \\ d_- \ \text{ with probability } \ 1 \ - \ p_{\text{H}} \end{cases}$$

We assume that the two possible values of D_{Conf} satisfy $d_+ > d_{\text{CW}}$ and $d_{\text{C}-} < d_{\text{CW}}$. In comparison to the model of Figure 3a, Figure 3b treats all positive values of D_{Conf} as "lumped" together at d_+ , and all negative values of D_{Conf} as "lumped" together at d_- . Within this model, we say that Defender is of type *Hard* if $D_{\text{Conf}} = d_+$, and that Defender is of type *Soft* if $D_{\text{Conf}} = d_-$.

In the binary model (Figure 3b), Defender is Hard if and only if Defender prefers *Conflict* to *Challenger Wins*, i.e., if and only if Defender prefers to execute its assumed threat to Resist in response to Initiation (see Figure 2). By analogy, we make the same definition for the continuous model (Figure 3a): We say that Defender is Hard if $D_{\text{Conf}} > d_{\text{CW}}$ and Defender is Soft if $D_{\text{Conf}} < d_{\text{CW}}$.[7] In general, we say that a Defender is Hard if and only if it prefers to shift the game unilaterally to *Conflict,* rather than accept the *Challenger Wins* outcome.

---

[5] The probability distribution of a random variable that can take any value in a continuum, such as an interval of real numbers, can be described by a *probability density function*, or *pdf*. To interpret the pdf in this case, suppose that an interval on the horizontal axis of Figure 3a has been specified. The probability that D_{Conf} falls into this interval equals the area of the region that lies over that interval but below the curve of the pdf. Note that, for any pdf, the total area over the horizontal axis and under the curve of the pdf must equal 1.

[6] A binary random variable can have only two possible (numerical) values, in contrast to a continuous random variable, which can have infinitely many. In the model of Figure 3b, the binary random variable D_{Conf} must equal either d_+ or d_-.

[7] Throughout, we will ignore transitional cases, such as the possibility that $D_{\text{Conf}} = d_{\text{CW}}$ in the model of Figure 3a. For a continuous random variable such as D_{Conf} in Figure 3a, this is an "event of probability zero," because the region under the pdf where $D_{\text{Conf}} = d_{\text{CW}}$ is true has area zero.

In fact, the probability that Defender is Hard should be related to Defender's credibility, the left side of (*), that refers to the perceived likelihood that Defender will execute its threat to respond to initiation. In general, we define p_H to be the probability that Defender is Hard, i.e.,

$$p_H = \Pr\{\text{Def is Hard}\} = \Pr\{D_{\text{Conf}} > d_{\text{CW}}\}.$$

Note that we have already used the symbol p_H in the model of Figure 3b, but, as is easy to verify, the above definition coincides with that usage. In the continuous model (Figure 3a), p_H is the area of the shaded region to the right of d_{CW} in Figure 3a. To complete the analysis of the Rudimentary Asymmetric Deterrence Game, note that we have already determined the optimal strategy for Challenger, which is Not Initiate if

$$p_H \geq \frac{c_{\text{CW}} - c_{\text{SQ}}}{c_{\text{CW}} - c_{\text{Conf}}},$$

and Initiate otherwise. As stated above, this relation is identical to (*), provided we interpret p_H as Defender's credibility, $c_{\text{CW}} - c_{\text{SQ}}$ as Challenger's gain at *Challenger Wins*, and $c_{\text{SQ}} - c_{\text{Conf}}$ as Challenger's cost of *Conflict*. It is easy to verify that an optimal strategy for Defender is to plan to Resist if Hard, and to plan to Not Resist if Soft. The Rudimentary Asymmetric Deterrence Game always (except in transitional cases) has a unique equilibrium, which is deterrence if and only if the deterrence effectiveness condition holds. Otherwise deterrence fails, and the outcome is *Conflict* if Defender is Hard, and *Challenger Wins* if Defender is Soft.

To summarize, the Rudimentary Asymmetric Deterrence Game is particularly simple model of a deterrence situation. Deterrence is effective if and only if the deterrence effectiveness condition (*) holds. Moreover, this game provides two useful representations of Defender's value for Conflict; understanding variations in this quantity is crucial to understanding the effects of force downsizing

73

Models of the Effects of Force Downsizing

Our analysis of the Rudimentary Asymmetric Deterrence Game as a game of one-sided incomplete information does not provide us with any way to model the perceived change in cost of conflict for Challenger, as we have not tried to "unpack" Challenger's values.[8] But the two models in Figure 3 do give us a picture of how downsizing would affect Defender's credibility.

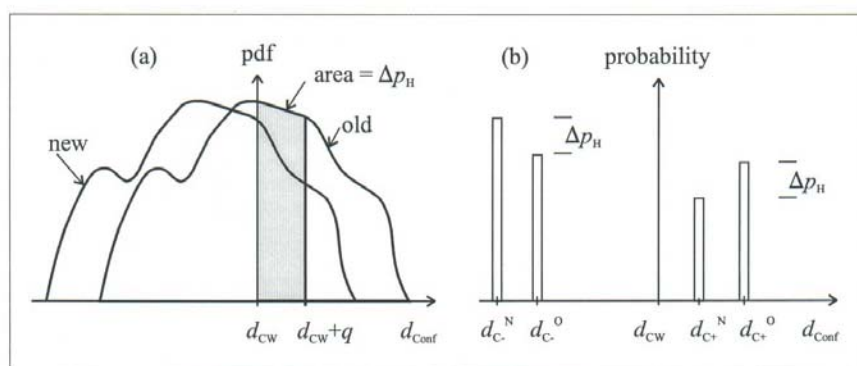The effects in models 3a and 3b are depicted in Figures 4a and 4b, respectively.

Figure 4: Models of the Effect of Downsizing on Credibility

Figure 4a shows a model of the effects of force downsizing in the context of the continuous model of Figure 3a. Note that the entire distribution of Defender's values has been shifted to the left by an amount q. But all that is relevant to deterrence effectiveness is the difference in p_H. This difference equals the probably that, in the original distribution, Defender's value for conflict, D_{Conf}, satisfied

$$d_{CW} < D_{Conf} < d_{CW} + q .$$

---

[8] In Kilgour and Zagare (1994a), we explore how components of Challenger's and Defender's evaluation of the issues at stake affect the stability of an extended deterrence relationship. Kilgour (1991) applied essentially the same idea to a different deterrence model.

74

The probability of this event equals the shaded area in Figure 4a that lies between d_{CW} and $d_{CW} + q$. If the distribution of D_{Conf} is the roughly "bell-shaped," as shown in Figure 4a, then this probability may be unexpectedly large relative to the value of q, since the interval concerned is central, and therefore captures a relatively large amount of probability.

Figure 4b shows a model of the effects of Defender's force reductions in the binary model of Figure 3b. Figure 4b shows that Defender's utility is typically reduced—whether it fights or not; in addition, Defender may be less likely to fight. The only change relevant to (*) is that Defender's credibility has been reduced by Δp_H. The meaning of this quantity is shown graphically in Figure 4b. Note that the change in deterrence effectiveness depends on Δp_H only. Thus, these models of Defender's value for conflict can connect models of downsizing of forces to the success of deterrence. One important illustration is the all-or-nothing (massive retaliation) deployment policy that, as noted above, seems to describe the direction of recent developments in international security.

"All-or-Nothing" Deployment Policies

An all-or-nothing deployment policy is, essentially, the threat to retaliate against initiation using a response of enormous destructive power. This is a way of downsizing forces that has been claimed to enhance deterrence effectiveness. The idea is that Defender builds a cheaper response system that actually increases the cost of conflict to Challenger. For instance, in the 1950s, many strategic analysts in the United States believed that atomic and nuclear weapons provided "more bang for the buck" and were both less expensive and more effective than large US installations and conventional troop commitments to Europe. The "New Look" policy of the Eisenhower administration and the associated doctrine of massive retaliation was a policy to reduce the resource consumption of US forces. Similarly, the British deployment plan prior to World War I was essentially an all-or-nothing approach, as was France's prior to World War II. Before 1914, Britain relied almost exclusively on an escalatory threat (i.e., its fleet) to

75

deter German expansion. And during the inter-war years, French deterrence policy rested on the threat of "massive firepower" to ward off aggression. Increasingly, the Russian deployment is an all-or-nothing approach.

Some simple but insightful views of all-or-nothing deployments can be derived from the models presented above. One important aspect of an all-or-nothing stance can be depicted using a binary model that is a variant of Figure 4b. In this variant, the right-hand spike (or probability mass) moves far to the right, indicating that Defender's response will be very powerful, if it is made at all. However the right-hand spike continues to be associated with a reduced probability; as already noted, this probability is p_H, and it is the value of p_H, that determines, via (*), deterrence effectiveness. In other words, what makes deterrence work is the probability associated with the right-hand spike. Reductions in this probability cannot be compensated by moving this spike further to the right.

An all-or-nothing deployment policy causes changes in Defender's value for *Conflict*, which can be seen as causing changes in Defender's credibility. Figure 4 can be used to trace these effects. Another (obvious) consequence of an all-or-nothing deployment is a reduction in Challenger's value for *Conflict*. The effects of such a change can be seen in Figure 1. There, the position after an all-or-nothing policy has been adopted must lie to the right of where it began, rather than to the left. Figure 1 suggests that this effect of an all-or-nothing deployment might be consistent with continued deterrence success—provided reductions in credibility are compensated by sufficient increases in the cost of *Conflict* to Challenger.

To resolve these questions requires a more detailed and specific study. We carried out such a study in Zagare and Kilgour (1993b), and concluded that policies like massive retaliation are, for strategic reasons, generally ineffective. There will generally be some lower-level conflicts that cannot be deterred by the threat of a disproportionate response (that both sides would disprefer), and lower-level conflicts can sometimes escalate to higher levels. Moreover, these facts depend only on higher levels of conflict being more costly to both sides, and are

independent of the credibility question: Who would believe a threatened response that is far out of proportion to the provocation? In other words, even when both players have highly credible endgame or strategic-level threats, massive retaliation is not an effective deployment policy.

Thus, we believe that a state that sidesteps the fighting effectiveness issue by shifting potential destruction to a higher level pays a high price in deterrence effectiveness. Even if willingness to intervene is unchanged, strategic considerations mean that the threat to fight is not as likely to deter unwanted actions. Thus, an important component of the value of the forces is lost.

Conclusions

Is downsizing a good idea? We have demonstrated some ways to conceive of the costs of downsizing in terms of threat effectiveness. Of course, these ideas have not led to any practical schemes for measuring this effectiveness. Nonetheless, we believe that we have made an important contribution to understanding the balance between cost savings through downsizing of forces and strategic losses due to the weakening of those forces' ability to deter.

Our conclusions, drawn from a variety of models under many informational and credibility assumptions, are consistent with the recent recommendations of the National Academy of Sciences. Since our models predict threshold effects, we support reductions in strategic weapons to the point at which they are clearly effective, but would no longer be clearly effective if reduced further. In our models, more is not necessarily better. We do not claim any special insight into the precise number of warheads that are at once sufficient and effective, but the Academy's recommendation of 200 land-based and 100 submarine-based warheads strikes us as plausible.

On the other hand, a total reliance on all-or-nothing threats is rarely effective. Our models reveal that highly credible conventional and related sub-strategic forces are also required to deter low-level conflicts. In our examination of limited war deployment policies, we find that

deterrence is most effective when both tactical and strategic credibility is high (Zagare and Kilgour, 1995). Significantly, no commitment to escalate first is required to support the conventional deterrence equilibrium that emerges when Defender has relatively credible tactical and strategic-level threats. Thus, our models also support the National Academy of Sciences recommendation of a no-first-use policy, provided sub-strategic forces are maintained at a significant level. Otherwise, conventional force reductions could undermine strategic deterrence.

References

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

Curtis, Willie, 2000. "The Assured Vulnerability Paradigm: Can it Provide a Useful Basis for Deterrence in a World of Strategic Multi-Polarity?" *Defense Analysis*, 16: 239 – 56.

Danilovic, Vesna, 2001a. "Conceptual and Selection Bias Issues in Deterrence." *Journal of Conflict Resolution*, 45: 97 –125.

Danilovic, Vesna, 2001b. "The Sources of Threat Credibility in Extended Deterrence." *Journal of Conflict Resolution*, 45: 34 – 69.

Gacek, Christopher, 1994. *The Logic of Force: The Dilemma of Limited War in American Foreign Policy*. New York: Columbia University Press.

Gaddis, John Lewis, 1997. *We Now Know: Rethinking Cold War History*. New York: Oxford University Press.

Geller, Daniel S. and J. David Singer, 1998. *Nations at War: A Scientific Study of International Conflict*. Cambridge: Cambridge University Press.

Gholz, Eugene and Harvey M. Sapolsky, 1999/2000. "Restructuring the U.S. Defense Industry." *International Security*, 24: 5 – 51.

79

Harvey, Frank P., 1998. "Rigor Mortis, or Rigor, More Tests: Necessity, Sufficiency, and Deterrence Logic." *International Studies Quarterly*, 42: 675 – 707.

Huth, Paul K., 1999. "Deterrence and International Conflict: Empirical Findings and Theoretical Debates." *Annual Review of Political Science*, 2: 61 · 84.

Intriligator, Michael D. and Dagobert L. Brito, 1984. "Can Arms Races Lead to the Outbreak of War?" *Journal of Conflict Resolution*, 28: 63 – 84.

Kegley, Charles W. and Eugene Wittkopf, eds., 1989. *The Nuclear Reader: Strategy, Weapons, War*. 2nd ed. New York: St. Martin's.

Kilgour, D. Marc, 1991. "Domestic Political Structure and War Behavior." *Journal of Conflict Resolution*, 35: 266 – 84.

Kilgour, D. Marc and Frank C. Zagare, 1991. "Credibility, Uncertainty, and Deterrence." *American Journal of Political Science*, 35: 303 – 34.

Kilgour, D. Marc and Frank C. Zagare, 1994a. "Uncertainty and the Role of the Pawn in Extended Deterrence." *Synthese*, 100: 379 – 417.

Kilgour D. Marc and Frank C. Zagare, 1994b. "Can Escalation Be Controlled? A Game-Theoretic Analysis." Paper delivered at the 35th Annual Meeting of the International Studies Association, Washington, DC, March 29.

Levy, Jack S., 1988. "When Do Deterrent Threats Work?" *British Journal of Political Science*, 18: 485 – 512.

National Academy of Sciences, 1997. *The Future of U.S. Nuclear Weapons Policy*. Washington, DC: National Academy Press.

O'Neill, Barry, 1992 "Are Game Models of Deterrence Biased Towards Arms-Building?" *Journal of Theoretical Politics*, 4: 459・77.

Paul, T.V., Richard J. Harknett, and James J. Wirtz, eds., 1998. *The Absolute Weapon Revisited*. Ann Arbor: University of Michigan Press.

Quinlan, Michael, 2000/2001. "How Robust Is India-Pakistan Deterrence?" *Survival*: 42: 141 – 54.

Sanger, David E. and Erick Eckholm, 1999. "Will Beijing's Nuclear Arsenal Stay Small or Will It Mushroom?" *New York Times*, March 15.

Tammen, Ronald L. et al., 2000. *Power Transitions: Strategies for the 21st Century*. New York: Chatham House.

Turner, Stansfield, 1999. *Caging The Genies: A Workable Solution for Nuclear, Chemical, and Biological Weapons*. Boulder, CO: Westview Press.

U.S. Department of Defense, 1998. *Annual Report to the President and the Congress*.

Washington, D.C.: Department of Defense.

Zagare Frank C., 1996. "Classical Deterrence Theory: A Critical Assessment," *International Interactions*, 21: 365 – 387.

Zagare, Frank C. and D. Marc Kilgour, 1993a. "Asymmetric Deterrence." *International Studies Quarterly*, 37: 1 – 27.

Zagare, Frank C. and D. Marc Kilgour, 1993b. "Modeling 'Massive Retaliation'." *Conflict Management and Peace Science*, 13: 61 – 86.

Zagare, Frank C. and D. Marc Kilgour, 1995. "Assessing Competing Defense Postures: The Strategic Implications of 'Flexible Response'." *World Politics*, 47: 373 – 417.

Zagare, Frank C. and D. Marc Kilgour, 1998. "Deterrence Theory and the Spiral Model Revisited." *Journal of Theoretical Politics*, 10: 66 · 87.

Zagare, Frank C. and D. Marc Kilgour, 2000. *Perfect Deterrence*. Cambridge: Cambridge University Press.