

**FORCE AND RESTRAINT IN STRATEGIC
DETERRENCE:
A GAME-THEORIST'S PERSPECTIVE**

Roger B. Myerson

November 2007

Visit our website for other free publication downloads

<http://www.StrategicStudiesInstitute.army.mil/>

[To rate this publication click here.](#)

This publication is a work of the U.S. Government as defined in Title 17, United States Code, Section 101. As such, it is in the public domain, and under the provisions of Title 17, United States Code, Section 105, it may not be copyrighted.

The views expressed in this report are those of the author and do not necessarily reflect the official policy or position of the Department of the Army, the Department of Defense, or the U.S. Government. This report is cleared for public release; distribution is unlimited.

This paper is based on a talk presented at the Chicago Humanities Festival on Peace and War, November 11, 2006.

Comments pertaining to this report are invited and should be forwarded to: Director, Strategic Studies Institute, U.S. Army War College, 122 Forbes Ave, Carlisle, PA 17013-5244.

All Strategic Studies Institute (SSI) publications are available on the SSI homepage for electronic dissemination. Hard copies of this report also may be ordered from our homepage. SSI's homepage address is: *www.StrategicStudiesInstitute.army.mil*.

The Strategic Studies Institute publishes a monthly e-mail newsletter to update the national security community on the research of our analysts, recent and forthcoming publications, and upcoming conferences sponsored by the Institute. Each newsletter also provides a strategic commentary by one of our research analysts. If you are interested in receiving this newsletter, please subscribe on our homepage at *www.StrategicStudiesInstitute.army.mil/newsletter/*.

ISBN 1-58487-325-6

FOREWORD

This monograph is a short nontechnical introduction to the use of game theory in the study of international relations. The focus is on the problem of deterrence against potential adversaries and aggressors. The author, Professor Roger Myerson, uses game models to provide a simple context where we can see more clearly the essential logic of strategic deterrence. We should look to such theoretical analysis for basic insights that may have practical importance in policymaking.

The main conclusion is that a great power's use of its military forces may be rendered ineffective or even counterproductive when there are no clear internationally recognizable limits on this use of force. Professor Myerson derives this conclusion from the basic observation that our ability to influence potential rivals depends on a balanced mix of threats and promises. Potential adversaries should believe that aggression will be punished, but such threats will be useless unless they also believe our promises that good behavior will be better rewarded. A reputation for resolve makes threats credible, but a great power also needs a reputation for restraint, to make the promises credible as well. Thus, international restraints on a nation's use of military force may actually increase the effective influence of its military strength. So this monograph may be read as a contribution to our understanding of the vital relationship between diplomacy and military preparedness in defense of national security.



DOUGLAS C. LOVELACE, JR.

Director

Strategic Studies Institute

BIOGRAPHICAL SKETCH OF THE AUTHOR

ROGER B. MYERSON, winner of the 2007 Nobel Prize in Economics, is the Glen A. Lloyd Distinguished Service Professor of Economics at the University of Chicago. He taught for 25 years in the Kellogg School of Management at Northwestern University before going to the University of Chicago in 2001. Dr. Myerson is the author of two books and many professional articles on game theory, information economics, and economic analysis of political institutions. He is a Fellow of the American Academy of Arts and Sciences and has served as its Midwest Vice President. Dr. Myerson was elected Vice President of the Econometric Society in 2006. He won a Guggenheim Fellowship in 1983 and received an honorary doctorate from the University of Basel in 2002. Dr. Myerson holds a Ph.D. from Harvard University.

SUMMARY

In a dangerous world, we need to think very carefully about how military force is used. Game theory can serve us in such analyses by providing a framework for probing the inextricable connections between our adversaries' decision problems and our own. To illustrate the power of game theory, the author focuses on a vital question that confronts American policymakers today: What determines why an application of military force, which was intended to deter potential adversaries, sometimes instead stimulates them to more militant reactions against us? When we feel that force is necessary, what can we do to minimize the risk of such adverse reactions?

A successful deterrent strategy is key and requires a balance between resolve and restraint, and this balance must be recognized and understood by our adversaries. So for our forceful actions to have their intended deterrent effect, they should be framed by a process of communication with our potential adversaries that establishes mutually recognized limits and rules about what we will and will not do.

From early roots in the work of John von Neumann and John Nash, game theory developed as a general framework for analyzing systems of incentives that involve two or more rational actors. Applications of game theory have extended beyond the traditional scope of economics to include the design of auctions, incentives in organizations, analysis of political institutions, and problems of international relations. In game-theoretic analysis of international relations, the great seminal classic is Thomas Schelling's *Strategy of Conflict* (Harvard University Press, 1960). In particular,

the vital importance of our strategic coordination with our adversaries, as well as with our friends, was shown by Schelling and is a fundamental point of this paper. Indeed, all arguments herein may be viewed as straightforward applications or extensions of Schelling's ideas.

FORCE AND RESTRAINT IN STRATEGIC DETERRENCE: A GAME-THEORIST'S PERSPECTIVE

In a dangerous world, we need to think very carefully about how military force is used. Game theory can serve us in such analyses by providing a framework for probing the inextricable connections between our adversaries' decision problems and our own. To illustrate the power of game theory, I focus here on a vital question that confronts American policymakers today: What determines why an application of military force, which was intended to deter potential adversaries, sometimes instead stimulates them to more militant reactions against us? When we feel that force is necessary, what can we do to minimize the risk of such adverse reactions?

A successful deterrent strategy requires a balance between resolve and restraint, and this balance must be recognized and understood by our adversaries. So for our forceful actions to have their intended deterrent effect, they should be framed by a process of communication with our potential adversaries that establishes mutually recognized limits and rules about what we will and will not do.

From early roots in the work of John von Neumann (1928)¹ and John Nash (1950)², game theory developed as a general framework for analyzing systems of incentives that involve two or more rational actors. Applications of game theory have extended beyond the traditional scope of economics to include the design of auctions, incentives in organizations, analysis of political institutions, and problems of international relations. In game-theoretic analysis of international relations, the great seminal classic is Thomas Schelling's

Strategy of Conflict (Harvard University Press, 1960). In particular, the vital importance of our strategic coordination with our adversaries, as well as with our friends, was shown by Schelling and is a fundamental point of this paper. Indeed, all arguments herein may be viewed as straightforward applications or extensions of Schelling's ideas.

Intuitive Arguments Against Multilateral Constraints on American Actions.

Before making the case for strategic restraint, let me begin by citing some important expressions of a contrary view, against the acceptance of multilateral constraints on America's use of force. For example, consider the comments of President George W. Bush in September 2002, when a reporter asked his opinion about Democratic senators who did not want to consider the option of invading Iraq without a prior decision by the United Nations (UN). The President responded:

Democrats waiting for the U.N. to act? I can't imagine an elected member of the United States Senate or House of Representatives saying, "I think I'm going to wait for the United Nations to make a decision." It seems like to me that if you're representing the United States, you ought to be making a decision on what's best for the United States. If I were running for office, I'm not sure how I'd explain to the American people: "Say, vote for me, and, oh, by the way, on a matter of national security, I think I'm going to wait for somebody else to act."³

The logic of the President's view seems clear. To defend ourselves against foreign threats, we may sometimes need to use military force against our adversaries, and any external constraint might prevent us from using such force when we feel it is

necessary. To protect American national security, American leaders have a responsibility to make these vital decisions about American military actions, and it would seem best for them to choose from the broadest possible range of military options according to their own best judgment. Accepting multilateral constraints that reduce the scope of American military options would seem counterproductive, like destroying part of our own military hardware, which was acquired at great expense to expand the capabilities of our forces.

In the President's view, this argument against accepting foreign constraints on American military actions seems so clear and so universal that the question of accepting UN restraint is not even worth discussing in an electoral campaign. He feels sure that any attempt to defend a policy of accepting multilateral constraints on American military actions would be a losing political strategy.

The President here was speaking off the cuff. It might be worthwhile to consider also a careful articulate expression of the argument against multilateralism by Jonah Goldberg in 2006:

If it was right to topple Saddam Hussein, it was right even if no one else agreed. If it was wrong, then it was wrong even if the world was on our side. Lynch mobs aren't right because they have numbers on their side, and men who stand up to them aren't wrong because they stand alone. Multilateralism is good only to the extent that it allows us to achieve good things.⁴

The image of the lone defender of justice is a good one (although we might be a bit disturbed if our local police chief actually applied this argument to rely on his own best judgment in arresting people, regardless of what any judge or jury would say). Goldberg acknowledges

that we might want UN approval if we needed their help to conquer Iraq in the first place. But if we can beat Saddam Hussein on our own, he argues, then we should ask no other question than whether it would be good or bad for us to do so.

This argument implicitly assumes that an invasion is intrinsically right or wrong, regardless of what others think. But a fundamental goal of our military strategy is to deter others from aggression against our country. So the long-run success of a deterrent strategy for protecting America depends on how foreigners throughout the world will respond to our actions. Thus, if we care how others will react in the future, then we may indeed want our decisions to take account of their judgments. This is the basic insight that I will try to develop in this paper, using game theory.

In such questions of deterrence, where the best strategy for us depends on how others will react to it, our strategic plan should be based on careful analysis of the actions that our potential adversaries will choose. But when we seriously endeavor to understand the choices of our adversaries, we may realize that their best plan of action must be based on their analysis of how we are likely to react to them. So we cannot understand our decision problem or our adversaries' unless we analyze our decisions and theirs together as part of an inextricably connected whole. Game theory has been developed as a framework for analyzing such interconnected decision problems.

Game Theory as a Form of Analytical Narrative.

Game theorists study mathematical models of social interactions. To be useful, a game model should be simple enough to understand but should share some important similarities with the more complex

situations of conflict and cooperation that we face in real life. Game theorists use models as simplified versions of life that are meant to clarify some of the logic of life's dilemmas, just as people everywhere use stories to develop new perspectives on important social problems.

People regularly tell stories to help themselves to understand society and its problems. To understand an international crisis, we might seek useful analogies by retelling, for example, the story of the 1938 Munich appeasement (to justify resolve) or the 1962 Cuban missile crisis (to justify restraint). In every culture, people have accumulated a wide repertoire of stories that are regularly retold as analogies that guide people's thinking about social problems. Of course, no story that we can tell will fit any real situation perfectly. To overcome the limitations of one story, we need to consider many stories, and the insights that we get from different stories must be compared. Models in game theory are just stories of another kind.

As any literary form, game theory has stylistic constraints which define both the power and the limitations of game theory. The people in the game are the players, and each player in the game has to choose an action from a given set of alternatives. In the game, players have goals which are described numerically by payoffs that depend on everybody's actions. It is assumed that each player wants to maximize his own expected payoff. Players may learn some information in the game, and we can describe this possible information in a game model by listing the set of possible observations with numerical probabilities for each.

So game theory requires a mathematically precise description of what each person can do in the game and

the preferences are that guide each person's choice of action. Such precision has the advantage that it makes game situations very clear, but it requires us to eschew the subtle vocabulary and imagery that enriches other forms of story-telling.

In the analysis of such game models, game theorists always try to respect the players. In particular, we assume that the players are intelligent, in the sense that they understand everything that we game-theorists understand about their game; and we assume that each player is rational, in the sense that he will always choose his own action to maximize his own expected payoff. In game-theoretic analysis, an *equilibrium* (as defined by John Nash) is a prediction of all players' actions such that each player's action is best for himself, given what the other players are predicted to do. Nash equilibrium is our basic solution concept for understanding what people can rationally do in a game.

A Simple Model of the Strategic Deterrence Problem.

To describe a dangerous world where incentives for aggression are pervasive, let us consider a version of the well-known "Prisoners' Dilemma" game, shown in Table 1 below. This game involves two players, whom we name simply "A" and "B." For interpretation, let us say that player A represents America, and player B represents some other smaller country in the world. In this basic game, each player must simultaneously choose one of two possible actions: cooperation or aggression. The payoff for each player depends on both of their actions, as shown in the table. For each pair of actions, Table 1 lists two numbers, the first being A's payoff and the second being B's payoff.

	B cooperative	B aggressive
A cooperative	0, 0	-8, 1*
A aggressive	*1, -8	* -3, -3*
	A's payoff, B's payoff	

Table 1. A Game with Pervasive Incentives for Aggression (The Prisoners' Dilemma).

The asterisks indicate the best payoff that each player could get in response to each possible action of the other player. If B were expected to be cooperative, then player A could get the payoffs 0 or 1 by cooperation or aggression, and so the best response for A is indicated by the asterisk before A's payoff 1 in the bottom-left cell. On the other hand, if B were aggressive, then player A could get the payoffs -8 or -3, and the best response for A in this case is indicated by the asterisk before A's payoff -3 in the bottom-right cell. The best responses for B to each of A's possible actions are similarly indicated by asterisks after B's payoffs, in the top-right and bottom-right cells here. The cell that has two asterisks is a Nash equilibrium of the game, because here each player is choosing his best response. In Table 1 we see that the unique equilibrium of this game is in the bottom right, where both players are aggressive.

In the simple structure of this Prisoners' Dilemma game, each player finds his aggressive action to be his best response to each of the other player's possible actions, but each player's choice of aggression rather than cooperation is very harmful to the other player. So both players get payoff -3 in the unique equilibrium of the game. Of course, they would both be better off with payoff 0 if both would cooperate, but mutual cooperation is not an equilibrium, as each player

will always be tempted to aggression. The root of the dilemma here is that each player always gets at least a small gain by switching from cooperation to aggression himself, but the result of such aggression will be a large loss for the other player. So when the players have no opportunity to respond to each other's actions in this game, each player wishes that the other player would act cooperatively but knows that he has no incentive to do so.

The analysis would change, however, if one of the players could observe the other's action first and respond to it. If A's decision to act cooperatively or aggressively could depend on what B chooses to do, then A's reaction could give B some positive incentive to cooperate. In this context, we can talk meaningfully about deterrent strategies.

A *strategy* for a player in a game is a complete plan that specifies an action for the player in every possible situation that the player could encounter in the game. Modifying the game of Table 1, let us now suppose that player A can observe whether B chooses cooperation or aggression before A makes his own choice between cooperation and aggression. When player A gets to move second after observing what B does, player A has four possible strategies which are listed in Table 2.

A's strategy: B's action:	B cooperative	B aggressive
A is cooperative always	A cooperative	A cooperative
A does the same as B	A cooperative	A aggressive
A does opposite of B	A aggressive	A cooperative
A is aggressive always	A aggressive	A aggressive

Table 2. The Four Strategies for Determining A's Action when A Can Observe B's Prior Action.

Now the outcome of the game will depend on B's choice and A's strategy as shown in Table 3 below. In each cell, the payoffs are those from Table 1 when player A chooses the action that is specified by his strategy against the given action of B. For example, when A's strategy is "do the same as B," if B is cooperative then A is cooperative and the resulting payoffs (from Table 1) are 0 for each; but if B is aggressive, then A is aggressive, and the resulting payoffs (again from Table 1) are -3 for each.

	B cooperative	B aggressive
A is cooperative always	0,0	-8, 1*
A does the same as B	0,0*	* -3, -3
A does the opposite of B	*1, -8	-8, 1*
A is aggressive always	*1, -8	* -3, -3*
	A's payoff, B's payoff	

Table 3. A Game Where Player A Moves after Observing B's Action.

In each row of Table 3, an asterisk after the second number in a cell indicates that it is the best payoff that B can get in response to the strategy of player A in this row. Notice that a best-response asterisk appears after B's payoff in the left B-cooperative column only for one strategy, the strategy where A does the same as B. So player A here has one *deterrent strategy* that motivates B to act cooperatively, and that is the strategy where A does the same as B.

But is this deterrent strategy credible? When B is cooperative, player A would get payoff 0 by doing the same as B, but player A could get the higher payoff 1 by using a strategy that is aggressive in this case. So A does not want to actually follow his deterrent

strategy when B cooperates, so B should not believe that A would use this deterrent strategy, unless A can somehow constrain himself to follow this strategy. Without such restraint, this game still has only one equilibrium, where both players are aggressive and both get -3 .

So to induce B to cooperate, player A wants to make a credible commitment to follow the deterrent ("do same as B") strategy, but this commitment requires some outside force to restrain player A from acting aggressively when B has cooperated. Thus, player A would prefer to enlarge this game by adding some other players who could punish A for acting aggressively when B has cooperated. Such punishment can actually be achieved if A is expected to play similar games in the future, if the behavior of future opponents in subsequent games can depend on how A behaves in this game now.

To be specific, let us suppose that player A will play a game like the one in Table 3 every year but with a different player "B" each time. Suppose that player A has a reputation for using the "do same as B" strategy in these games, against which the B players should act cooperatively so that A's payoff should be 0 in every game. But if player A ever lost that reputation by acting aggressively against a cooperative B-player, then we may suppose that the mutual-aggression equilibrium would be played in all future games, yielding the payoff -3 in all future games. At a 5 percent annual interest rate, a income stream that pays \$3 every year would be worth \$60 in present discounted value (because depositing \$60 in a bank account that pays 5 percent annual interest would allow you to take \$3 income every year forever). So A's reputation in this repeated game should have a present value R that is

approximately $R=60$. Subtracting this lost reputational value from A's payoff in the cases where A has acted aggressively against a cooperative opponent, the current game against the current player B looks like Table 4. As long as the reputational value R is greater than 1, there is a good equilibrium in which B is cooperative and A does the same as B.

	B cooperative	B aggressive
A is cooperative always	*0,0	-8, 1*
A does the same as B	*0,0*	* -3, -3
A does the opposite of B	1-R, -8	-8, 1*
A is aggressive always	1-R, -8	* -3, -3*
	A's payoff, B's payoff	

Table 4. A's Reputation for Restraint Worth $R>1$ Can Make Deterrence Credible.

The game in Table 4 also has a bad equilibrium where everybody is always aggressive, and both players get the bad payoff -3. Indeed, the threat of switching to such a bad equilibrium in the future is what sustains the good equilibrium in Table 4. So with multiple equilibria, communication between the players may help to get everyone focused on the better equilibrium, according to Schelling's focal-point effect. That is, to coordinate everyone's expectations on the good reputational equilibrium, player A might announce to the world:

We promise to be aggressive only when we have prior proof of our opponent's aggression. If we ever deviate from this promise, then you may expect us instead to be aggressive always, so that all our opponents should be aggressive against us if we break this promise even once. But as long as we maintain our reputation, our

opponents should expect to gain by cooperating. So you should be confident that we will rationally maintain our reputation for restraint, because getting \$0 always is better for us than getting \$1 once and then \$-3 always afterwards.

Thus A can benefit from cultivating a reputation for restraint in eyes of the world.

For this negotiation speech to be persuasive, however, everyone must anticipate that they would not be persuaded by it again after A was seen deviating from the terms of the promised strategy. If A could take aggressive profits (1) against one cooperative B-player and then persuade the next B player that they should resume the good reputational equilibrium where the Bs all cooperate with A, then A would want to be aggressive every time. So the other future B players in this repeated game all need to understand the promised terms of A's deterrent strategy, and they need to actively monitor A's behavior and judge whether A has acted correctly according to this strategy.

Judging Reputations.

We have been assuming that when a player is aggressive, the whole world will see that player's aggression. Let us consider what happens when this assumption is dropped in the game in Table 4, where player A has a reputation for restraint that has a large long-run value R to player A. Suppose now that, if B is aggressive, the whole world will probably see it, but there is some small positive probability ϵ that only player A will see B's aggression and everyone else will think that B has been cooperative. What should A do in a game where such a disagreement about B occurs? When the world thinks that B has been cooperative,

it is better for A to be cooperative now and preserve the reputation than to be aggressive now and lose the reputation ($-8 > -3$ B R). As long as the probability of B's aggression escaping general detection is not too large ($\epsilon < 3/4$ so that $\epsilon \times 1 + (1-\epsilon) \times -3 < 0$), the threat of A's aggressive response when the world community recognizes B's aggression should be sufficient to deter such aggression.

On the other hand, the credibility of the deterrent strategy could not be sustained if A were to act aggressively on private evidence that nobody else can observe. If A would not lose any reputation by being aggressive when the world sees no evidence of B's aggression, then A would prefer to be aggressive always, claiming always to be justified by private evidence of B's prior aggression; and so the deterrent strategy would not be credible, as in Table 3. Thus, A's reputation must be judged by others.

This important point deserves some emphasis. We have argued that a nation's military actions must be judged as part of a deterrent strategy, and this judgment cannot be made by the nation itself when it has any ability to benefit from such actions. In such situations, the jury that passes judgment on a nation's military actions must be outside the nation itself.

Reputations for Restraint and Resolve.

The example in Table 4 is intended as a simple model of American foreign policy, with player A representing America. In this model, we have a deterrent strategy that involves both a promise that we will cooperate if our current rival cooperates, and a threat that we will be aggressive if our current rival is aggressive. For our deterrent strategy to be effective,

our rivals must believe our promise of cooperation. But in the repeated Prisoners' Dilemma game, we could always get a short-run benefit from acting aggressively instead of cooperating. So to make our promises of cooperation credible, we need some reputational commitment to act cooperatively when our deterrent strategy promises it, and such a commitment is what we mean by *restraint*. So this model illustrates how the credibility and effectiveness of America's deterrent strategy may require us to maintain a reputation for accepting restraint from others in the UN.

Of course, an effective deterrent strategy also requires that our rivals must believe our threats of punishment as well as our promises of cooperation. In other situations where fighting is costly and aggression is unprofitable, we might also need some reputational commitment to act aggressively when our deterrent strategy threatens it, and such a commitment to act aggressively in such situations is what we mean by *resolve*. So a reputation for resolve may also be needed to make credible the threats that an effective deterrent strategy also requires. Such a reputation for resolve would be lost if we were ever seen to cooperate with a rival who had acted aggressively against us. Then the fear of losing such a reputation could give us an incentive to respond aggressively when we would otherwise find it unprofitable.

Which is more important: resolve or restraint? In the repeated Prisoners' Dilemma game analyzed here, credible deterrence only required a reputation for restraint. A reputation for resolve was unnecessary in that game, because short-run incentives could always motivate aggressive action. Of course, the Prisoners' Dilemma is just one simple model. More complicated models can be formulated in which effective deterrence

requires a reputation for resolve as well as a reputation for restraint. But remember why the Prisoners' Dilemma interested us in the first place: because the problem of deterring aggression becomes acute only to the extent that individuals gain short-run profits from aggressive actions that hurt others. If nobody ever had a problem of credible restraint from aggressive behavior, then we would not have to worry about deterring aggression in the first place. So there is good reason to believe that restraint should be a significant part of most real deterrence problems.

One could argue, however, that resolve might be more important than restraint for small weak nations, because their weakness makes conflict more dangerous for them. But stronger nations can find more opportunities for profitable aggression, and so they may have less need for resolve and correspondingly more need for restraint. By this argument, we should expect restraint to be most important for America, when America is the world's most powerful nation.

For a simple example where resolve may be more important, consider the game in Table 5 (which differs from Table 1 in that the -3 and -8 payoffs have been switched). Now each player's best response is to be aggressive if the other player is cooperative (as $1 > 0$), but to be cooperative if the other player is aggressive (as $-3 > -8$). So the paired asterisks in the bottom-left cell indicate an equilibrium where player A is aggressive and player B is cooperative, which is the best possible outcome for A but is bad for B. But there is also another equilibrium in the top-right cell where player A is cooperative and player B is aggressive, which is the best possible outcome for B but is bad for A.

	B cooperative	B aggressive
A cooperative	0,0	*-3, 1*
A aggressive	* 1, -3*	-8, -8
	A's payoff, B's payoff	

Table 5. Game with Multiple Equilibria Where a Reputation for Resolve Could Be Valuable.

For such games with multiple equilibria, Schelling (1960) argued that anything in the players' shared culture or environment that focuses their attention on one equilibrium can lead them to act according to it, as a self-fulfilling prophecy. In particular, if player A has a reputation for being aggressive in such games, then B may naturally focus on the equilibrium that player A prefers. Furthermore, if player A would lose this valuable reputation by cooperating with an aggressive player B, then this reputation for resolve could transform the game into one where the unique equilibrium is the outcome that player A most prefers, as shown in Table 6. But of course, if both players try to maintain a reputational commitment to acting aggressively in such games, then they would both suffer the worst payoff -8.

	B cooperative	B aggressive
A cooperative	0,0	-3 -R, 1*
A aggressive	* 1, -3*	* -8, -8
	A's payoff, B's payoff	

Table 6. A's reputation for resolve worth $R > 5$ makes B cooperative in equilibrium.

On the other hand, Table 7 shows another variation on Table 5 in which player A is stronger than B, but A's strength is counterproductive because it effectively stimulates B's resolve. In this game, the strong player A has an extra option to invade the other nation. If A invades when B is aggressive, then the outcome is the same as if A were merely aggressive; but if A invades when B is cooperative, then A enjoys a higher payoff of 2 while B suffers a lower payoff of -9. Player B would still be willing to cooperate if player A were simply aggressive, but A's temptation to turn aggression into invasion here eliminates any equilibrium where B is cooperative. So player A's extra option in Table 7 makes this a game with a unique equilibrium, which coincides with the worst equilibrium for player A in Table 5. So the strong player A here is only harmed by his ability to invade and could actually benefit by an external constraint that would eliminate this option for him.

	B cooperative	B aggressive
A cooperative	0, 0	*-3, 1*
A aggressive	1, -3*	-8, -8
A invades	*2, -9	-8, -8*
	A's payoff, B's payoff	

Table 7. The Perceived Danger of Being Invaded by A Makes B Aggressive in Equilibrium.

Failure of Deterrence: Stimulating Militarism by Denying Restraint.

More generally, we should ask, why might rational citizens prefer militant leaders who are committed to costly military actions? People could have two natural

motivations for such militarization: to profitably attack others, or to decrease their own chances of being attacked. Most small countries in the world have few opportunities for profitable aggression, but a desire for stronger defenses against possible attacks can be a motivation for people to seek militant leadership in any country. Militant leaders can demonstrate the resolve that smaller countries especially need when they fear a possibility of invasion, as in Table 7. So to decrease our neighbors' incentives to militarize against us, we want to reduce their fears of being attacked. Our forceful acts without clear strategic limits can counterproductively increase unconquered adversaries' militant commitment against us.

President Bush announced in September 2001 that nations anywhere in the world that support terrorism will be treated by the United States as hostile regimes.⁵ This proclamation may have sounded like a strong deterrent strategy, but its effective implementation would depend critically on who judges whether a nation is or is not supporting terrorism. In the comments quoted above, the President insisted that America's leaders should make these judgments themselves, with no multilateral constraints. Thus, there should be serious questions about whether this broadly aggressive policy could fail as a deterrent strategy because of a lack of clear restraint. Our demonstrated willingness to preemptively invade a country on our own private interpretation of evidence, with no external constraints on our use of force, could be seen as a dangerous repudiation of strategic restraint, which could inspire counterforces against us (from guerrilla to nuclear, depending on local capabilities).

In particular, suppose that people in some nation had some reason to think that America might want

to attack them. Then they would naturally fear that America's judgment about whether they support terrorism could be swayed by American interests against them, regardless of whether they have actually supported terrorism or not. Indeed, they might well see America's 2003 invasion of Iraq as hard evidence of the possibility that they might also be so invaded. But if they believe that such an American invasion is likely, no matter what they do, then they might rationally calculate that their security could actually be enhanced by sponsoring global terrorism, to keep more American forces busy elsewhere in the world. They might also view their development of nuclear weapons as another way to improve their own security, by making the contemplated invasion much riskier for America. Thus, America's refusal to accept multilateral restraint could actually exacerbate terrorist insurgencies and nuclear weapons proliferation.

So our theoretical models suggest that support for global terrorism today might actually have been less if the Bush administration had not conspicuously rejected UN restraint earlier in this decade. Of course, it is impossible to prove this hypothesis as we cannot relive this decade with a different policy decision, nor can we compare this decade to an otherwise identical period when great-power leaders showed more willingness to accept external restraint. So these fundamental questions about the effectiveness of our deterrent strategies must be guided by logical analysis, and our conclusions may depend on the assumptions that we bring to our model. Other models with other assumptions might lead to different conclusions. On such important policy questions, we need a debate in which different models and views are compared.

But in September 2002, the President's expressed opinion was that accepting UN restraint on America's

military decisionmaking would be such a bad idea that no responsible politician should even suggest it, if he hopes to get reelected. As we have seen from our models, however, there is good reason to think that a reputation for accepting UN restraint may have been exactly what America needed to deter terrorist aggression in this decade. The logic of our analysis here should be sufficient to make the case, at least, that the question of accepting UN restraint was worthy of serious political debate in America. It should not have been mockingly dismissed by someone with a responsibility to think about what is best for America.

For another recent example of a deterrent strategy executed without sufficient clarity of limits or restraint, we may consider the Israeli retaliation in Lebanon during the summer 2006. Israeli bombing was widespread and lacked any clearly articulated limits or proportionality to the provocation. There was no doubt that the kidnapping of soldiers was an aggressive provocation that required some strategic response; but the heavy and unfocused nature of Israel's response could raise questions about whether its intention was to not merely deter specific acts of aggression but also to change the nature of the regime in Lebanon to Israel's benefit. Whatever the true intentions were, such questions could readily stimulate Lebanese fears of a deep invasion that would renew the violence of their civil war. A natural response to such fears of invasion would be to support militant parties who seem more capable of fighting Israel. Thus, in the aftermath of the 2006 summer war, the leader of Hezbollah, Hasan Nasrullah, could preside over a huge rally in Beirut where he posed as Lebanon's strongest defender, calling 1200 Lebanese deaths in the recent war a small price to avoid Iraq's fate of "10,000 to 15,000 people killed every month in a chaotic war incited by the

Americans and the Mosad.”⁶ Thus, ambiguity about the limits of American and Israeli military actions may have helped Hezbollah to sell itself as a strong defender of Lebanon’s security.

Conversely, Arab calls for total elimination of Israel motivate Israelis to bear the high costs of their militarization. People everywhere want security against any perceived risk of a devastating invasion. Retaliatory actions and threats that lack clearly defined limits can raise fears of deep invasions and thus can motivate people on the other side to seek militant leadership that may be better able to defend them. Lack of clear restraint can stimulate others’ resolve.

Thus, if we want our application of military force to deter our potential adversaries, rather than stimulate them to more militant reactions against us, then we should make sure that the limits of our forceful actions are clear to any potential adversaries. We need a reputation for responding forcefully against aggression, but we also need a reputation for restraining our responses within clear limits that depend in a generally recognized way on the nature of the provocation. These limits must be clear to our potential adversaries, who must be able to verify that we are adhering to the limits of our deterrent strategy, because it is they whom we are trying to influence and deter.

Reconsidering the Rationality Assumption.

Game-theoretic analysis is based on an assumption that people are rational. Of course nobody is perfectly rational; we all make mistakes. But to get a sense of what people are likely to do in any given situation, it is generally a good idea to think about what their interests are and to assume that they will act to pursue

these interests. Our adversaries' interests may be different from our own, but we generally share at least some common interests, such as avoiding the costs of destructive conflict.

But it may be asked: What if our adversaries are irrational or congenital aggressors who cannot be deterred? If so, what could we do but try to bind them or destroy them? We should be very cautious about jumping to such conclusions. After all, if our adversaries understood that we believed this about them, so that our perceived self-interest would require their destruction, then their struggle against us actually could become rational self-defense for them. Thus we should not lightly contemplate such self-fulfilling prophecies of congenital violence and mortal struggle. It is generally much safer to assume that our adversaries will respond appropriately to a firm deterrent strategy when our resolve and restraint are both made clear to them.

Of course, there are people in the world who are irrationally or pathologically drawn to violence and destruction. Our most dangerous adversaries are not lone madmen, however, but are leaders with political support from many people who have normal hopes and fears. Psychopathic militarists like Hitler become a threat to our civilization only when ordinary rational people become motivated to support them as leaders.

One might also question the game-theoretic assumption of selfish rationality when it is applied to our own country. That is, we may ask: What if the assumption of selfish rationality does not apply to us because we can always be trusted to do what is right? If so, then our intrinsic justice could be a sufficient deterrent against aggression, and we would not need to worry about maintaining our reputation for appropriate

restraint. But the effectiveness of our intrinsic justice as an implicit source of restraint depends critically on others believing in it. Our own belief in our own justice is not enough.

Lessons for Patriots.

This paper has considered simple game-theoretic models to probe the basic logic of deterrence. The lessons of this analysis may be briefly summarized.

To influence and deter potential rivals, we need a deterrent strategy in which threats of conflict are balanced by promises of cooperation when rivals yield to our pressure. The threats and promises of our deterrent strategy can be effective only if they are understood and believed by our potential rivals. When Americans judge our leaders for effectiveness in foreign policy, the central question should be how our policy is perceived by the foreigners whom we want to influence and deter. Letting these foreigners judge our reputation for adhering to our deterrent strategy can help us to guarantee its credibility. So a policy of submitting American military actions to international judgment and restraint can actually make America more secure.

Any bomb by itself can only cause destruction. Our bombing can have a constructive purpose only as part of a strategy that defines when we bomb and when we do not bomb. But people have no incentive to yield if they think that their yielding would only invite further aggression and invasion. If our rivals do not understand the limits to our use of military force, then our bombing can only spread destruction and resolve to resist us. So we want our adversaries to understand the limits on our use of force, and we want them to verify that we are complying with these limits.

Thus, we can benefit from a reputation for accepting strategic limits on our use of force, using proportionate retaliatory force only as necessary for deterring attacks on us. But strategic communication is essential. Our strategic limits must be clearly communicated to potential adversaries, because our reputation for resolve and restraint is effective only to the extent that our adversaries recognize it. For questions of whether our use of force has been appropriate under our deterrent strategy, the ultimate judge and jury are the potential rivals whom we want to deter and reassure. That is, on a question of whether a particular use of military force was justified under our deterrent strategy, the judgment that counts is that of the foreigners whom we want to deter. If our invasion has been justified only to American voters, then it has not been justified at all.

It has sometimes been suggested that Americans who doubt that everybody can always trust America should be denigrated as unpatriotic. Of course, patriots should have some basic faith in the good qualities of our country, but that does not imply that patriots must always assume that everybody in the world will accept our good faith without any proof or guarantee. As has been argued above, ignoring foreign fears of our power and denying any need for restraint can inadvertently stimulate more militant reactions against us in the world. So Americans who want to accept multilateral restraints, to reassure foreigners about the limits of our power, may be true patriots who are asking the hard questions that are essential to our national security.

ENDNOTES

1. John von Neumann, "Zur Theorie der Gesellschaftspiele" (On the Theory of Parlor Games), *Mathematische Annalen* 1928.
2. John Forbes Nash, Jr., "Equilibrium Points in N-Person Games," *Proceedings of the National Academy of Sciences*, Vol. 36, 1950, pp. 48-49.
3. President Bush discusses Iraq with reporters, September 13, 2003, at www.whitehouse.gov/news/releases/2002/09/20020913.html.
4. Jonah Goldberg, "When Multilateralism Falls Short," *National Review*, October 13, 2006.
5. In the President's address to Congress and the American people on September 20, 2001, at www.whitehouse.gov/news/releases/2001/09/20010920-8.html.
6. From Hasan Nasrullah's speech at a Hezbollah rally in Beirut, September 22, 2006, at www.juancole.com/2006/09/nasrullahs-speech-here-is-bbc-world.html.