FULLY INFORMED AND ON THE ROAD TO RUIN:

THE PERFECT FAILURE OF ASYMMETRIC DETERRENCE

by

Jean-Pierre P. Langlois*

Department of Mathematics, San Francisco State University

langlois@math.sfsu.edu

and

Catherine C. Langlois

McDonough School of Business, Georgetown University

Langlois@msb.edu

August 12, 2003

Revised September 12, 2004

Revised January 30, 2005

*The listing order of the authors' names is not indicative of their respective contributions, which they consider to be equal. We wish to thank an anonymous reviewer for valuable feedback and corrections.

ABSTRACT

Most theoretical and formal arguments about rational deterrence assume that war is a game-ending move. In the asymmetric case, the logic of deterrent threats then rests on the relative merits of war and submission. Perfectly informed rivals ensure that immediate deterrence always succeeds although general deterrence may not. Does this strong result survive the repetition of the standard one-shot deterrence game? We show that an unbundling of the war outcome, and the resulting possible recurrence of a challenge to the *status quo*, changes the very nature of deterrent threats and can lead to the failure of immediate deterrence. If the *status quo* can be challenged repeatedly, it is rational, in case of challenge, for the rivals to threaten probabilistic escalation of the crisis to war with the following consequences: the challenger will challenge the *status quo* now and then; the defender finds it rational to resist at least for a while; the resulting recurrence of challenge, resistance, and escalation can lead the rivals to threaten, with some likelihood, wars that are long enough to be catastrophic for all parties.

INTRODUCTION

Most theoretical and formal arguments about rational deterrence assume that war is a game-ending move. The logic of deterrent threats then rests on the relative merits of war and submission, and perfectly informed rivals will never fight, although the defender may fail to deter a challenger from upsetting the *status quo* if she cannot credibly threaten war.[1] Perfectly informed rivals therefore ensure that immediate deterrence always succeeds although general deterrence may not. Does this strong result survive the repetition of the standard one-shot deterrence game? We show in this paper that an unbundling of the war outcome, and the resulting possible recurrence of a challenge to the *status quo*, changes the very nature of deterrent threats and can lead to the failure of immediate deterrence. If the *status quo* can be challenged repeatedly, it is rational, in case of challenge, for the rivals to threaten escalation of the crisis to war with some probability. Equilibrium play therefore "keep(s) the enemy guessing" (Schelling, 1960:200) about the timing and duration of war. But while Schelling's "threat that leaves something to chance" appeals to the irrational slippage of a state that loses control, our fully informed rivals wield the probabilistic threat of war *strategically*.

Our approach is closely related to the work of Zagare and Kilgour (1993, 1998, 2000) and Zagare (2004) in its focus on asymmetric deterrence and the conditions for credible general and immediate deterrence. But Zagare and Kilgour's conclusions about general deterrence rely on games with under-specified "conflict" endgames. As the authors themselves point out, "game theoretic models are, in essence, empty vessels: they can be filled with a wide variety of substantive fluids," (Zagare and Kilgour, 2000:71). In particular, the "conflict" outcome could mean that the protagonists fight a brief and limited decisive war, or it could mean that they engage in a protracted battle for territory or influence. Conflict could mean Operation Desert Storm's forty-three days of war, or it could mean the hundred months of deadly battle between Iraq and Iran between 1979 and 1990. But war could also be the recurring outcome of a pattern of challenge and escalated threats, interrupted by long periods of unresolved latency, as in Peru's long standing border dispute with Ecuador. By allowing our rivals to engage in several rounds of conflict, back away from altercation, return to the *status quo*, submit, or repeatedly challenge, we also allow them to generate the variety of conflict outcomes that are observed in real-world settings. Importantly, we find that these outcomes also emerge from rational play. But if the repetition of the standard asymmetric deterrence game allows for the formal representation of protracted conflicts, it also upsets the calculus of perfect deterrence.

Our emphasis on asymmetric deterrence means that we unbundle the conflict outcome in the case where a clear defender enjoys a prize that is coveted by the challenger. As Zagare and Kilgour point out, such "one-sided deterrence relationships have obvious empirical and theoretical import" (Zagare and Kilgour, 2000:135). However, in our model, the rivals switch roles if the challenger wins. Actual possession of the prize is what determines a player's role in a game that repeats indefinitely, and each party gets a chance to deter the other from challenging or escalating the conflict when endowed with the contested asset. As a result of the unbundling of the conflict node, we identify a vast class of rational strategies, in the form of subgame perfect equilibria (SPE), with the following properties:[2]

1) The challenger will challenge the *status quo* with some probability;

2) The defender finds it rational to resist at least for a while;

3) The resulting recurrence of challenge, resistance, and escalation can lead the rivals to engage in a conflict that lasts so long that its cost outweighs any benefit the winner may enjoy once the dispute is finally over. In other words, the rivals threaten, with some likelihood, wars that are long enough to be catastrophic for all parties.[3]

The very possibility of ruinous warfare in equilibrium has behavioral implications. If the result of equilibrium play can be catastrophic, some attention should be paid to the impact of the rivals' choice of strategy on the likelihood of such outcomes. The vast class of SPEs described above involves probabilistic moves, with an array of possible equilibrium choices for the challenger. The challenger can forgo challenge altogether, ensuring the success of general deterrence, but he can also challenge with varying degrees of assertiveness that find measure in the chosen probability of challenge at the *status quo*. However, given the defender's equilibrium response to challenge, the defender finds himself *indifferent* between accepting the *status quo* and challenging it. Thus the defender's deterrent threat of escalation *can* lead to general deterrence success. But if general deterrence fails, the defender's response seals the probabilistic failure of immediate deterrence. So how should the challenger behave? He could challenge aggressively, implement a timid policy of infrequent challenge, or forgo challenge altogether. All options are equal from the standard *ex ante* discounted utility perspective. Yet, the various strategic choices available to the challenger determine very different futures. An aggressive stance may lead the defender to hand over the prize sooner, but it could also lead both parties to accumulate war costs in excess of the value of the prize. A balanced evaluation of these realizations can inform the challenger's choice of strategy in equilibrium.

We begin with a brief literature review and follow with a conceptual discussion of asymmetric deterrence in a one-shot game, focusing in particular on the interpretation of the end game. We then unbundle the final nodes of the one-shot game to allow for a process of possibly repeated challenge and protracted conflict, examine its equilibria and discuss the implications of this unbundling in the formulation of deterrent threats. Finally, we discuss our equilibria in light of their consequences. As we discuss strategy and outcomes, we highlight the parallels between our formal approach and the history of real world conflicts.

A BRIEF LITERATURE REVIEW

Our work finds root in two strains of the vast literature on war and peace: firmly set in the literature on rational deterrence, it also uncovers a possible strategic driver of enduring rivalry. Our subject is rational deterrence in a context where the rivals can repeatedly challenge the *status quo* and escalate the crisis. We find that, in equilibrium, the protagonists can adopt strategies that, implemented, would generate behavioral patterns observed between enduring rivals. As such, our work offers a purely strategic motive for recurrent armed conflict, although the modeling of enduring rivalries is not our central purpose. In a rare attempt to model enduring rivalries game theoretically, Maoz and Mor (1999, 2002) describe rivals that change their preferences over future game outcomes as they learn about the opponent's capabilities.[4] Recurring conflict results from the rivals' changing satisfaction with the current *status quo,* and the rivalry comes to an end when the protagonists' moves confirm their beliefs about relative capabilities. To generate their enduring rivalries, Maoz and Mor (1999, 2002) examine the non-myopic equilibria of a matrix game (Brams, 1994). As a result their model cannot support a discussion of rational deterrence and the credibility of threats since, as Zagare and Kilgour (2000) forcefully argue, rational deterrence requires that the players implement subgame perfect equilibrium strategies. By contrast, we show that the very credibility of threats can lead rational rivals to fight occasionally while the conflict festers unresolved for much of the time, a pattern that is characteristic of enduring rivalries.

As Lemke and Reed (2001) and Sartori (2003) point out, "the theory of enduring rivalries is as yet poorly developed," (Sartori, 2003:20). But much work has been done to uncover the causes of rivalry as well as those that lead rivals to fight. While definitional details vary, the empirical importance of enduring rivalries is well established. Gochman and Maoz (1984) find that over half of all disputes between 1816 and 1976 involve rivals that have engaged in conflict with each other more than seven times. Importantly, enduring rivals are also more likely to fight (see Goertz and Diehl, 1993 and 1998, among others). Domestic and systemic shocks, a territorial dispute, and power parity are frequently found to be empirically significant precursors to war among enduring rivals, although Lemke and Reed (2001) offer contradictory evidence. Huth and Russett (1993) and Huth (1996) find that relative capabilities impact the likelihood of war once a military threat has become manifest. Heldt (1999), drawing upon diversionary theories of war in which leaders fight to divert attention from domestic issues that could cost them their jobs (Downs and Rocke (1994), Smith (1996)), finds that domestic dissatisfaction increases the likelihood that states involved in a territorial dispute will use force. Vasquez (1996) finds that the dyadic war over territory is one of two empirically relevant paths to war for enduring rivals.[5]

Our model, which assumes that a defender holds and enjoys a prize that is coveted by the challenger, mirrors the stakes involved in a territorial dispute. As such, it can be viewed as uncovering a strategic path to recurrent dyadic fighting over territory. But it also impacts the possible interpretation of data on enduring rivalries. For example, Huth and Russett (1993), in a direct attempt to link deterrence to enduring rivalry data, interpret intervals between militarized conflict as periods of general deterrence success. A war episode is then an isolated incidence of deterrence failure, rather than the manifestation of an overarching deterrence strategy that requires occasional escalation of the conflict to war for the very threat of costly fighting to be credible. Our model predicts possibly long periods of unresolved latency during which the rivals avoid confrontation while keeping the threat of escalation alive. But intervals between fights do not signal general deterrence success in our model. Rather, their length and quality reflect strategic decisions on the probability of immediate deterrence failure.

While the model we develop captures some aspects of enduring rivalry, our primary goal remains a discussion of rational deterrence in a context where the war outcome is unbundled. Powell (2003) notes that "most formal studies of the causes of war treat the decision to go to war as a game-ending move."[6] However, our effort to unbundle the war outcome is not unique, although our discussion of rational deterrence in this context is an innovation. Indeed, following Wagner (2000), and motivated by his claim that the treatment of war as a game-ending assumption "can only lead to misleading conclusions," a number of authors have unbundled the war outcome to allow for intra-war bargaining. These models typically focus on war as a source of information. In Filson and Werner (2002), the attacker is uncertain about the defender's military capability and learns from the outcome of wars fought. The rivals in Smith and Stam (2001) update their beliefs about the likelihood of winning a fort as they battle for these forts, one at a time.[7] Powell (2003) models war as a costly process during which states can bargain while running the risk of military collapse if they fight. The model assumes that one state is uncertain about the other's war costs and likelihood of military collapse.

An interest in intra-war bargaining has been the primary motivation for the unbundling of the war outcome. And, in light of the widespread belief that rivals that are fully informed should always settle their differences peacefully in equilibrium (Powell, 1990 and Fearon, 1995), most of these models assume incomplete information.[8] Yet enduring rivals, engaged in disputes that can last for decades, should surely get to know each other. Our analysis assumes that the rivals are fully informed about each other's priorities and capabilities, thus ruling out imperfect information as the cause for deterrence failure. Few authors have attempted to explain war between fully informed rational actors. Slantchev (2003) and Garfinkel and Skaperdas (2000) are notable exceptions. Garfinkel and Skaperdas (2000) construct a two-period model of resource allocation in which each party builds an arms stock and decides on settlement or war. But arms built in the first period are assumed to decay dramatically if not used, and war is a game-ending move. Under these conditions Garfinkel and Skaperdas (2000) find that fully informed rivals rationally go to war. Slantchev (2003) describes fully informed rivals that can bargain or fight and identifies SPEs in which the rivals agree to a settlement that is delayed by a few turns of war. In Slantchev's equilibria the rivals *agree* to go through several turns of war under the threat of reversion to extremal equilibria. Since war does not mark the end of the game, Slantchev unbundles the war outcome.

Slantchev's objective is to demonstrate the existence of inefficient equilibria despite the fact that the rivals are fully informed. It is not to discuss the nature of deterrent threats, which is our purpose here. Early theoretical discussions of deterrence often revolved around the special issue of nuclear deterrence as in Schelling (1960), Snyder (1964), and Jervis (1984). But broader approaches abound from the classical work of Morgan (1977), to Ordershook (1989), Wagner (1992), and Zagare (1990). The more formal analysis can be traced to Brams and Kilgour (1988), Langlois (1991), Morrow (1989), and Powell (1987, 1990), among others, and more recently to Zagare and Kilgour (2000) and Zagare (2004). O'Neill (1989, 1994) and Morrow (2000) also offer comprehensive reviews of the vast game theoretic literature on deterrence. These authors identify imperfect information as the source of immediate deterrence failure. But the recurring nature of conflict, while clearly recognized in the empirical literature, is never associated with the rational failure of deterrence. Yet we will show that fully informed rivals, engaged in protracted conflicts, threaten war probabilistically. Their willingness to actually fight some of the time is what makes the deterrent threat credible.

ONE SHOT ASYMMETRIC DETERRENCE

The game theoretic logic of deterrence hinges on the credibility of retaliatory threats in the face of an assault by the challenger. In the classic sequential asymmetric deterrence[9] game illustrated in Figure 1 below, the challenger has the first move and may wait, staying with the *status quo* (SQ), or challenge. The defender, in turn, at node 2, can resist if challenged or submit, ending the game at (SB). Faced with a resisting defender, the challenger, at node 3, can escalate or back down, leading the rivals to final outcomes war (WR) or backdown (BD). Payoffs are normalized as follows: at SQ, the challenger gets nothing (0) and the defender keeps the prize (1) while at SB the challenger gets the prize (1) and the defender loses it, possibly incurring a cost $-e \leq 0$. At BD the challenger incurs an "audience" cost of $-a \leq 0$ while the defender enjoys a benefit $b \geq 1$. And at WR challenger and defender pay the costs of conflict $-c < 0$ and $-d < 0$ respectively.

<< Figure 1 here >>

The challenger's preferred outcome is to see the defender submit to his challenge, while the defender ranks a backdown by the challenger ($b \geq 1$) at the top of the list. Escalation to war is less desirable than the *status quo* for each of the players. In this one-shot approach, the question of whether escalation to war is a credible threat lies in its value relative to submission for the defender, and backdown for the challenger. When the parties prefer war to backdown or submission, respectively ($c < a$ and $d < e$), deterrence prevails because the threat of war is credible.

The logic of rational deterrence if conflict is the worst outcome for the challenger ($c > a$) relies on a somewhat negative deterrence result. In this case, the defender should resist a challenge because the challenger will then back down. So, knowing that the defender will resist, and that backdown is the inevitable next rational step, the challenger should stay with the *status quo*. But, as Zagare and Kilgour (2000:142) point out, the *status quo* prevails because the challenger *cannot* deter the defender from resisting. Nevertheless, success of both general and immediate deterrence is the outcome. General deterrence fails in the asymmetric game of Figure 1 under one set of circumstances: the challenger can credibly threaten escalation ($c < a$) while the defender prefers to submit than to fight ($d > e$). In this case, immediate deterrence succeeds since the rivals do not fight, but general deterrence fails. Similar conclusions are reached if the basic asymmetric deterrence game is enhanced by adding finitely many layers of escalation, as discussed in Zagare and Kilgour (1998) and Zagare (2004).

In the one-shot framework the payoffs associated with all final outcomes should capture the expected long run value of that outcome. In particular, in a model of asymmetric deterrence, outcome SB should value the handing over of the prize from defender to challenger in perpetuity. The interpretation of the payoffs at WR is more complex, however. Costs $-c$ and $-d$ may represent the expected value of the possible future occurrences should war break out: perhaps it lasts for some time, is interrupted by a temporary backdown by the challenger, can be won or lost with some likelihood, and can be followed by a future in which the prize is secured or lost. The cost of fighting, the number of turns of war and the likelihood of winning then determine how costly it is to reach state WR. As endgames go, the substantive content of WR is under-specified in the one-shot game model.

Consider, for example, the long standing border dispute between Ecuador and Peru: it finds its roots in the creation of each of these states in 1830 and 1829, and it survives the 1942 Rio Protocol that was to delineate the border between the two states. Peru and Ecuador fought for four months before agreeing to the Rio Protocol. But the Protocol's demarcation was incomplete, and Ecuador rejected its validity, claiming extensive territory in the Amazon basin. In January 1981, Peru bombed Ecuadorean outposts at Paquisha, killing two and wounding twelve. The military phase of the Paquisha incident lasted seven days (Krieg, 1986). In January 1995 Ecuador and Peru battled in the Cenepa valley for thirty-four days, claiming as many as 600 lives (Weidner, 1996). And while the 1998 Brasilia Presidential Act, signed by both parties, resolves the Rio Protocol's border impasse, the "risk that either country will choose to use military force to achieve territorial objectives (...) is far from eliminated." (Simmons, 1999:21). How could all this information be absorbed in a generic "war" state WR? Clearly, such an effort would erase the dynamics of the conflict, rob the parties of their ability to decide when to fight, and rule out any strategic content to the course of events. Unbundling the events of end node WR allows for dynamic brinkmanship. The protagonists can decide how much they will fight if at all, and to fight again if peace does not bring agreement. But these possibilities can change the calculus of rational deterrence.

UNBUNDLING THE CONFLICT OUTCOME:
GAME STRUCTURE AND PLAYER OBJECTIVES

In order to unbundle the events that are implicitly contained in the one-shot conflict outcome, we repeat the game of Figure 1. If it is rational for the challenger to wait repeatedly, no crisis develops and general deterrence succeeds. But should the challenger choose to challenge the *status quo,* whatever the history of play, the following developments become possible:

• The defender could submit, handing over the prize to the challenger. This would mark the dawn of a new *status quo* in which challenger becomes defender, and defender becomes challenger. The rivals then choose strategy wearing a new hat.

• The defender could resist, forcing the challenger to decide between escalation to war and backdown. Neither decisions is final: If the challenger escalates the conflict, the rivals fight for one turn. Having incurred the cost of one round of fighting, the challenger can choose to return to the *status quo* or to challenge the defender once again, potentially risking a new turn of war if the defender resists. If the challenger backs down when the defender resists, the challenger incurs an audience cost, the defender gains from this temporary victory, and the rivals return to the *status quo.* It is then up to the challenger to challenge again or to wait. If the challenger waits, the defender reaps the rewards of possession for one turn. The challenger can wait repeatedly, letting the defender collect rent over and over again. But, at any time, the challenger can choose to challenge the *status quo* again.

Figure 2, below, is an iterated version[10] of the game of Figure 1, and it distinguishes between four states: the *status quo* (SQ), backdown by the challenger (BD), war (WR), and submission by the defender (SB). The decisions made by the protagonists inevitably lead them to one of the states of the game:

(1) SQ, the *status quo,* marks the beginning of the game. It is visited anytime the challenger waits at the previous turn;

(2) BD is visited whenever the challenger backs down (after a round of challenge and resist) at the previous turn;

(3) WR is visited whenever the challenger escalates (after a round of challenge and resist) at the previous turn.

(4) SB is reached if the defender submits after being challenged. SB is also the *status quo* of a new game in which the players switch roles. In this new game, our rivals are faced with the same opportunities and challenges. But now the prize has changed hands, and the challenger will take on the defender's role, while it is the once defender of the *status quo* that can challenge the new state of affairs.

Payoffs associated with each of the four states are indicated in Figure 2, below, with the challenger's payoff listed first:

<< Figure 2 here >>

When one of the three role preserving states (SQ, BD, and WR) is reached, payoffs are made for the current period only. By contrast, when SB is reached the payoff to each player depends on what each party expects the other to do at the outset of the new game. Thus, the new challenger's payoff becomes what the *defender* expected when still in possession of the asset at SQ, while the *new* defender's payoff corresponds to the current challenger's expectations at SQ. We will elaborate on this idea when discussing equilibria of the iterated game. To facilitate comparison of the payoffs in the one-shot game of

Figure 1 and the corresponding iterated game of Figure 2, we examine the payoffs of the iterated game in the special case where the defender, by submitting, gives up the contested asset forever. In this case, the payoff at SB is the discounted future value to the rivals of the defender's submission in perpetuity.

To adapt the payoff structure of the one-shot game to the iterated game, it is necessary to distinguish between per-period payoffs and accumulated future payoffs. Moreover, in the iterated game, future payoffs must be discounted. Let the discount factor be w with $0 < w < 1$. The payoff at SB is an accumulated sum of future payoffs while the payoffs at SQ, BD and WR are one-period payoffs. To make these payoffs fully comparable, we pre-multiply any current payoff by $(1 - w)$. So, listing the challenger first and with reference to the payoff structure of the one-shot game of Figure 1, a *one-period* wait by the challenger yields player payoffs $(1 - w) < 0, 1 >$. If the challenger never challenges and always chooses to wait, the players obtain the discounted value to infinity of per period payoffs $(1 - w) < 0, 1 >$ or

$$\sum_{t=0}^{\infty}(1 - w)w^t < 0, 1 > \ = \ < 0, 1 >$$

Similarly, a *perpetual* backdown would yield $< -a, b >$, and a *perpetual* war would yield $< -c, -d >$, while *each visit* to BD and WR yields payoffs $(1 - w) < -a, b >$, and $(1 - w) < -c, -d >$, respectively. By the same logic, if the protagonists expect *perpetual* submission at SB, they expect payoffs $< 1, -e >$, with per period payoffs at SB reading $(1 - w) < 1, -e >$. Given this expectation at SB, it is of interest to compare the predictions of the one-shot game to those of our iterated game.

In the one-shot game, general deterrence prevails when the defender can credibly threaten to escalate the conflict to war because war is less costly than submission: $d < e$. It is well known that repeating the equilibrium of the one-shot game provides a perfect equilibrium of the iterated game. So general deterrence can succeed in the iterated game as it does in the one-shot game if $d < e$. But what does this mean? In the iterated game, d is the cost of *perpetual* war to the defender while e is the cost of *perpetual* submission. General deterrence can then succeed if the defender prefers to fight *forever* than to give up the prize in perpetuity.[11] That a state would have such preferences seems unlikely and suggests ruling out this payoff structure in the iterated game of Figure 2. Indeed, in the logic of an iterated game, the parties can consider fighting for a limited time, and it is the cost of *temporary* warfare that is meaningfully compared to alternative outcomes. By accepting to fight, if challenged, the defender can hold on to the asset for longer, getting rent every time there is a lull in the hostilities. By fighting the challenger hopes to get the defender to hand over the prize by submitting. All in all, fighting could bring about a positive outcome for each rival. Even if the *per period* cost of war exceeds the *per period* cost of submission for the defender, $(1 - w)d > (1 - w)e$, a *limited* war can still be preferred to a long submission. Indeed, by accepting to fight for some time, or intermittently, the defender hopes for the challenger's occasional backdown with the subsequent return to the *status quo ante*, letting him enjoy possession of the prize. And it seems reasonable to assume that *perpetual* war is worse than surrendering the prize forever ($d > e$). This is the payoff assumption that we will make in what follows by setting $e = 0$ for simplicity, while $d > 0$.[12] With the payoff structure of the iterated game worked out, we will now be able to examine SPEs of the iterated game that give state SB richer strategic content.

But first, the rival's objective must be spelled out. Challenger and defender are standard expected discounted utility maximizers and choose strategy accordingly. To express the players' objectives, consider a sequence of player moves through the graph of Figure 2: Given any current decision node of the graph, such a sequence is valued according to the future payoff states visited. If S_t denotes the payoff state visited at turn t, then a payoff path is a sequence $\sigma = \{\mathrm{S}_1, \mathrm{S}_2, .., \mathrm{S}_t..\}$. Such a sequence could cycle within the graph of Figure 2, visiting states SQ, BD and WR indefinitely, or it could end with the players switching roles in state SB. Each particular payoff path is the result of a sequence of decisions made by the players. For example, path $\sigma = \{\text{BD,WR}, \text{SQ}, \text{SB}\}$ viewed from decision node SQ1, would result from the sequence of choices "challenge, resist, backdown, challenge, resist, escalate, wait, challenge, submit," with the players switching roles at SB.

Player i's discounted value for a payoff path σ is defined as

$$V_i^\sigma = U_i(\text{S}_1) + \omega U_i(\text{S}_2) + .. + \omega^{t-1} U_i(\text{S}_t) + .. = \sum_{t=1}^{\infty} \omega^{t-1} U_i(\text{S}_t) \quad (1)$$

For example, path $\sigma = \{\text{WR}, \text{SQ}, \text{SB}\}$ has discounted value for the defender:

$$V_D^\sigma = (1-\omega)b - \omega(1-\omega)d + \omega^2(1-\omega)1 + \omega^3(1-\omega)E_D(\text{SB})$$

where the defender's expected payoff at SB is the payoff she expects as challenger in the new *status quo.* The standard formulation of player i's expected utility, viewed from any *decision* node N, is then

$$E_i(\text{N}) = \sum_{\sigma} P(\sigma) V_i^\sigma \quad (2)$$

where the sum is taken over all possible paths σ following N, and $P(\sigma)$ is the probability that path σ will be traveled according to the players' strategies (see Fudenberg and Tirole, 1995, Chapter 5). Given objectives and payoffs for the iterated game of Figure 2, we now turn to a class of SPEs in which the defender implements a strategy that *could* lead to general deterrence success, but seals the probabilistic failure of immediate deterrence in case of challenge.

THE FAILURE OF IMMEDIATE DETERRENCE AND THE RATIONALITY OF WAR

The logic of asymmetric deterrence leads us to examine the case where one period of war is more costly to the defender than submitting for one period ($(1-\omega)d > 0$). But this does not mean that, in case of challenge by a credible challenger, the defender necessarily submits. In fact we analyze below a whole class of equilibria in which challenge is followed by resistance by the defender because the challenger may respond by backing down. Interestingly, occasional backdown by the challenger will be found rational whether he prefers to fight than to back down ($c < a$), or not ($a < c$). Of course, because repeating the equilibrium of the one-shot game is a SPE of the iterated game, it is rational for the challenger to stay with the *status quo* forever if $c > a$, and for the defender to submit immediately in case of challenge if $c < a$. But, in a whole class of equilibria, aggressive players do not conform to the behavioral patterns inherited from the one-shot game: The challenger can seek to force the defender to eventually submit even when war is costly and $a < c$. And the defender will not cooperate in her own submission although she knows that the challenger prefers to fight than to back down when she resists ($c < a$).

More precisely, a whole class of SPE involves finely calibrated threats and counter threats based on the probabilistic intentions to challenge, resist and escalate by challenger and defender. These equilibria have a structure that depends on the relationship between the challenger's audience cost of backdown and the cost he incurs by fighting for one period. We describe our class of equilibria with reference to the decision nodes of Figure 2 above, examining each of the rival's rational decisions as the conflict evolves.

The Challenger's Decision to Challenge and to Escalate to War

In general, player decisions at each turn could depend on the entire prior history of play. But, as Figure 2 illustrates, many different histories can lead the players to a given state, and once in that state the players always face the same set of possible decisions. It is therefore natural to investigate strategic behavior that depends only on the current "state of the game," regardless of any specific prior history. Such strategies are called Markov strategies because they yield a pattern of play akin to a Markov chain. A Markov perfect equilibrium (MPE) is simply an equilibrium in Markov strategies that holds at every state of the game. Focussing on Markov perfect equilibria has three main advantages: MPEs only require a specification of the players' intentions at each of their decision states and they are relatively easy to construct; MPEs are also SPEs, meaning that a player cannot benefit by deviating in any way from an MPE strategy; MPEs are in fact representative of a wide class of equilibria since "extremal equilibria" (the worst SPEs for each player) are usually constructed as MPEs.[13]

In the class of MPEs that we are interested in, the challenger behaves as follows:

At SQ1, the challenger challenges with probability s

$$\text{At BD1, the challenger challenges } \begin{cases} \text{with probability } t \text{ if } a \leq c \\ \\ \text{with probability 1 if } a \geq c \end{cases} \tag{3a}$$

$$\text{At WR1, the challenger challenges } \begin{cases} \text{with probability 1 if } a \leq c \\ \\ \text{with probability } t \text{ if } a \geq c \end{cases} \tag{3b}$$

At SQ1, the starting point of the game, and the point of return if the challenger decides to wait after waging war, our challenger can choose to wait or to challenge probabilistically. Probability s can be as small as the challenger wants, and it has an upper bound that depends on parameter values.[14] The challenger can therefore choose from a range of strategic options, from the timid to the tough. Our model therefore suggests that the initial challenge of a *status quo* can be a strategic matter rather than the result of serendipitous events that culminate to determine state action. China's military philosophy, with its emphasis on controlled offensive action (Johnston, 1995), could be an illustration of such strategic thinking. Syria's proposal to implement an "openly protracted struggle" to weaken Israel prior to the June 1967 war also suggests a strategically chosen frequency of challenge on the part of Arab states (Reiser, 1994:78).

The relationship between the challenger's per period audience cost of backdown $(1-\omega)a$ and the cost of war in each period $(1-\omega)c$ marks a critical break in rational behavior. A high audience cost of backdown $(a \geq c)$ moderates the challenger's bellicosity when war has broken out, but ensures that the *status quo* will be challenged for sure after a costly backdown.[15] By contrast, low audience costs relative to the cost of fighting $(a < c)$ prompt an aggressive strategy in war, with the challenger always challenging again after a round of fighting at WR1. But, after a backdown, the challenger can rationally adopt a flexible stance, deciding to challenge again infrequently by choosing a small t. However, it is also rational for the challenger to keep the defender under pressure at BD1 by picking a probability of challenge t as high as 1.

It is hard to evaluate a state's audience cost of backdown. However, the relatively low cost of the limited militarized disputes between Ecuador and Peru in the Amazon basin suggests that backdown may have been politically more costly than war for Ecuador. Ecuador's unilateral 1960 declaration that the Rio Protocol was null and void was confirmed at the outset of all the militarized confrontations with Peru until 1995, and many border incidents during the period reaffirmed Ecuador's ongoing challenge of the Rio Protocol. But in only three cases between 1960 and 1995 did the rivals escalate the conflict to wars that remained limited in time and in cost (Huth, 1996, Simmons, 1999). Ecuador's bellicosity was contained once war broke out, but especially after the Paquisha incident of 1981, challenge after backdown was persistent. This is the type of behavior that our model would suggest if audience costs to the challenger are high relative to the costs of war.

Table 1 below provides ranges for the challenger's choices at SQ1, BD1 and WR1 depending on parameter values. We set the following parameters: discount factor $\omega = 0.99,$ the defender's one-period benefit from challenger backdown $(1-\omega)b = 0.01 \times 2,$ and the defender's one-period cost of war $(1-\omega)d = 0.01 \times 2.$

We also examine the range of rational challenge behaviors for various values of parameters c and a:[16]

<< Table 1 about here >>

If the challenger's audience cost of backdown and war costs are similar, little restriction is put on probability s in equilibrium. As illustrated in Table 1, limits to the challenger's rational propensity to aggress at SQ1 arise when war is costly relative to backdown or vice versa. Explanation for this lies in the impact of the challenger's choice on the defender's decision to resist. As we will see below, the challenger does not need to challenge with high probability at SQ1 for the defender to submit at BD2 or WR2 with high probability when $\left|a-c\right|$ is large. The upper bounds on s, when they fall short of 1, represent hostility levels that lead the defender to give up, choosing to submit with certainty at BD2 or WR2. But as a result, the challenger will challenge with certainty at BD1 after a costly backdown (challenge at BD1 when $a=4$), but he can be more flexible if backdown is less costly than battle (challenge at BD1 when $a=1$). By contrast, having fought a costly war, the challenger challenges again with certainty at WR1 (challenge at WR1 when $a=1$), but remains more flexible if war is less costly than backdown (challenge at WR1 with probability t when $a=4$).

The challenger can choose from a wide array of strategies in equilibrium. In particular, he can always choose never to challenge, picking $s=0$. But he can also challenge the *status quo* to varying degrees depending on the history of conflict. His choices will determine the defender's response, and if he chooses to challenge at all, the defender will resist with some probability. In the face of a resisting defender, the challenger must anticipate a possible escalation of the conflict to war. At SQ3, BD3, and WR3, the challenger escalates with probability

$$q = q_1 = \frac{\omega(1-t)+b\big(1-(1-s)\omega\big)}{\omega(1-t)+(b+d)\big(1-(1-s)\omega\big)} \qquad \text{if } a \leq c \qquad (4a)$$

$$q = q_2 = \frac{b\big(1-(1-s)\omega\big)}{-\omega(1-t)+(b+d)\big(1-(1-s)\omega\big)} \qquad \text{if } a \geq c \qquad (4b)$$

q is therefore a probability that varies with the challenger's strategic choice of s and t.[17] Numerical values for q, are given below, setting the discount factor $\omega = 0.99$, $b = 2$, and $d = 2$ and illustrating the relationship between s, t and q setting parameters $a = 1$ and $c = 1.5$:

<< Table 2 about here >>

When on the brink of war, the challenger's long-term strategy, embodied in probabilities of challenge s and t, affect his decision to escalate when met with resistance by the defender. The more aggressive the challenger at the outset, the lower the probability of escalation to war when on the brink. Thus, as shown in Table 2 above, the challenger will escalate the conflict to war with 71% probability if he chooses to challenge infrequently in the first place. But he manages the risk of costly war by accepting to back down more frequently if, instead, he chooses to challenge the *status quo* more forcefully (q drops to 0.52 when $s = 1$ and $t = 0.8$).

The challenger's behavior in equilibrium is set in full knowledge of the defender's priorities and capabilities and, therefore, with full awareness of the defender's response. The defender holds the prize and enjoys the fruits of possession, and she will not be willing to hand it over on demand.

Deterrence and the Defender's Decision to Resist

The defender holds the prize and wants to keep it. A few turns of war may be a price worth paying if the challenger subsequently accepts the *status quo*, at least for some time, before challenging again. If challenged, the defender will rationally threaten to resist with the following probabilities:

$$\text{At SQ2, the defender resists } \begin{cases} \text{with probability } p_1 = \frac{1}{1+a(1-\omega(1-s))} & \text{if } a \le c \\ \\ \text{with probability } p_2 = \frac{1}{1+c(1-\omega(1-s))} & \text{if } a \ge c \end{cases} \quad (5a)$$

$$\text{At BD2, the defender resists} \begin{cases} \text{with probability } p_1 \text{ if } a \leq c \\ \text{with probability } r_2 = \frac{\omega + (c-a)\left(1-\omega(1-s)\right)}{\omega\left(1+c\left(1-\omega(1-s)\right)\right)} \text{ if } a \geq c \end{cases} \tag{5b}$$

$$\text{At WR2, the defender resists} \begin{cases} \text{with probability } r_1 = \frac{\omega + (a-c)\left(1-\omega(1-s)\right)}{\omega\left(1+a\left(1-\omega(1-s)\right)\right)} \text{ if } a \leq c \\ \text{with probability } p_2 \text{ if } a \geq c \end{cases} \tag{5c}$$

The defender resists with probabilities that depend on the challenger's choice of probability s. Once again setting $\omega = 0.99$, $b = 2$, and $d = 2$, Table 3 gives the defender's response to selected choices for s given parameters a and c:

<< Table 3 about here >>

The figures of Table 3 show the defender lowering her probability of resistance as the challenger becomes more aggressive by choice of s. This is a measured response to aggression that accounts for the likelihood of a possibly repeated escalation of the conflict to a war that is costly for both sides. Nevertheless, the defender calibrates her response according to the costs incurred by the challenger. If audience costs of backdown are high relative to the costs of war for the challenger ($a = 4,\ c = 1.5$), then the defender, anticipating the costs that can be imposed on the challenger, will resist more firmly after a round of costly war (at WR2), but will tone down her resistance after a backdown (at BD2). Symmetrically, when $a = 1$ and $c = 1.5$, comparison of the probabilities of resistance at BD2 and WR2 show the defender resisting more firmly after backdown than she does after a fight.

The nature of the defender's strategy is to make the challenger *indifferent* between all possible choices of s. In view of the defender's strategy, it is equivalent for the challenger to challenge at SQ1, or to simply accept the *status quo*. Thus, general deterrence *can* succeed as a result of the defender's choices. But if general deterrence fails, and the challenger decides to challenge, then immediate deterrence will also fail with some probability because the defender credibly threatens to resist despite the possibility of costly war. The logic of rational deterrence in the class of equilibria that we investigate here involves spelling out the *failure* of immediate deterrence if general deterrence fails. It is then up to the challenger to take the initiative, one way or the other. Such thinking is apparent among Arab and Israeli leaders prior to 1967. The credible threat of Israeli escalation in case of challenge was clearly understood by Egypt's President Nasser prior to June 1967 (Lieberman, 1995:869). And as Reiser (1994:87) reports, "He (Nasser) had resisted Syria's call for "incremental violence" against Israel...on the grounds that the Arabs would have no control over Israel's escalated level of retaliation." And Israel's willingness to escalate the conflict to war if challenged was clearly articulated in General Yigal Allon's policy of conventional deterrence: "The deterrence would be made up of astute political maneuverings and *an unknown but manageable number* of Israeli battlefield victories over an *unspecified but reasonable* period of time," (Reiser, 1994:81). Allon's anticipated "battlefield victories" do not determine an end to the conflict. Rather, they impose costs on the enemy and lead to backdown and temporary return to the *status quo*. The implementation of our rational strategies would also have such consequences.

The challenger's indifference between challenge and acceptance of the *status quo* at SQ1 sheds additional light on the figures of Table 3. The challenger's expected payoff from accepting the *status quo* forever is 0. But, given the defender's response in equilibrium, he also expects a 0 payoff when challenging with probability s at SQ1. The defender's calibrated threat of resistance makes war sufficiently likely for the expected benefits of challenge to match its expected costs for the challenger. The only possible benefit of a challenge is the defender's eventual submission. But what actually happens if the defender submits? In our model, the rivals switch roles, and this defines the expected payoff to the rivals in state SB. At SB, the defender hands over the prize but becomes the challenger of the new *status quo.* She therefore anticipates an expected payoff at SB of 0.[18] But what does the challenger expect if the defender indeed submits? At SB he expects to be the new defender and therefore anticipates the same expected payoff as the current defender at SQ. It turns out (see Lemma 1 in Appendix) that the defender's expected payoff at SQ is $\frac{1-\omega}{1-\omega+\omega s}$ and, in this class of equilibria, this is the expected payoff to the challenger at SB. In his new role as defender, the current challenger expects his rival to behave with the same level of aggressivity as measured by s, although the new challenger need not behave identically in her choice of probability t.

Our rivals can anticipate their changing roles, and their strategic choices are based on an evaluation of expected payoffs. But expected calculations mask the variety of outcomes that could result from an implementation of the rational strategies that we have described. Part of the story remains untold. In particular, the rational behaviors we discuss could, with some probability, lead the rivals to incur war costs that exceed the value of the prize. This is the result of the rival's probabilistic decisions in equilibrium. Indeed, equilibria in pure strategies could not threaten devastating wars for certain since such a threat would not be credible. This last point rejoins Schelling's discussion of "threats that leave something to chance," but while Schelling interprets the probability of devastating conflict as being exogenous or out of the decision maker's control, our rivals *choose* strategies that "keep the enemy guessing" (Schelling, 1960:200). Probabilistic moves in equilibrium can be interpreted as a state's ability to wield veiled threats whose credibility is the result of the possible consequences of a process that develops in time and can conceivably lead to long and protracted costly conflict. Schelling (1960:182), in his discussion of the randomization of promises and threats, writes, "it is interesting to notice that attaching a probability of fulfillment to our threat is....substantially equivalent to scaling...the size of the threat." The probabilistic moves chosen by our perfectly informed rivals serve the same purpose. The size of the threat is randomized strategically.

Our rivals are fully informed, yet they create uncertainty about the outcome of the crisis by choice of strategy. The probabilistic threat of war is one that leads the rivals to anticipate a range of possible war costs, each occurring with some likelihood. Nevertheless, the rivals could both end up enjoying a strictly positive payoff when roles are changed. This can happen if few wars are fought although the conflict remains unresolved for a long period of time. The defender then enjoys possession for sufficiently long, and the challenger eventually gets his turn in possession of the prize, having fought sparingly. But there are other possible outcomes. In particular, war costs could accumulate beyond the discounted value of the prize. This possibility and its likelihood is part of the defender's deterrent threat of resistance if challenged. But the consequences of a challenge depend on how aggressive it is. The challenger can choose to challenge with high or low frequency at SQ1, and this will be instrumental in determining the outcomes. While the defender's response makes him indifferent between all possible choices for s from a standard *ex ante* expected utility perspective, a more aggressive stance (higher s) will increase the likelihood that the accumulated costs of war will exceed the value of the prize. A higher likelihood of ruinous warfare is balanced by the likelihood that the defender will submit faster, handing over the valuable asset. The defender's choice of s is informed by an examination of the possible endgames that it could determine.

Exploring the Paths to War

Parameter values and choice of strategy determine the expected length of the crisis and the expected number of turns of war. We define the average frequency of war as the ratio of these two numbers. The dispute begins with the challenger's expressed dissatisfaction with the *status quo* and ends, if general deterrence fails, when the defender submits and the rivals switch roles. Within this possibly protracted dispute, the rivals can fight periodically, accumulating a total number of war turns. If the challenger is aggressive, choosing to challenge with high probability at SQ1, BD1 or WR1, the conflict will be shorter but more violent. Given strategy, higher war costs on both sides shortens the duration of the dispute but also reduces its violence. Table 4 below provides some selected data. We set parameters $d = 2,\ b = 2$ and $\omega = 0.99$ to examine the impact of changing war and audience costs for the challenger:

<< Table 4 about here >>

Table 4 compares dispute and war durations given a choice for s and t that is within all the allowable ranges determined by parameter values. Given $s = 0.3$ and $t = 0.8$, an increase in the challenger's per-period war costs $(1 - \omega)c$ shortens the crisis and decreases the frequency of in-crisis war. This is the result of the defender's response to renewed challenge after a turn of war. When c is high, the defender does not need to resist with as high a probability to impose a given cost on the challenger; comparing Cases 1 and 2, r decreases when c increases given $a = 1$, and comparing Cases 3 and 4, p decreases when c increases given $a = 4$. But this also means that when war costs to the challenger increase, the defender will submit with higher probability after fighting. Crisis length and overall frequency of war decline as a result. A comparison of Cases 1 and 3 and Cases 2 and 4 shows that an increase in the audience cost of backdown for the challenger also reduces crisis length and war frequency. This is the result of a subtle interaction between the challenger's probability of escalation q when faced with a resisting defender and the defender's probability of resisting after a backdown. The data in Table 4 shows that q changes little when audience costs increase, but the defender will submit with much higher probability after a backdown when audience costs are high.[19]

A more aggressive challenger determines a shorter dispute but a higher frequency of war given parameter values. For example, a choice of $s = 1$, $t = 1$ given $a = 1$ and $c = 1.5$ shortens the crisis to 1.79 periods but increases the frequency of war to 40%.

Nevertheless, our model predicts that expected war frequencies will remain contained. If faced with a challenge, the defender responds by threatening to resist probabilistically. This comes with a likelihood of escalation to war. But the defender calibrates her response so that the expected costs of possible warfare are compensated for by the expected benefits that she continues to receive as long as she does not give up the contested asset. At SQ, the defender's expected payoff is 0. In expected terms, the defender will not submit too soon, and war will not be waged too often. But these are expectations. Many possible outcomes lie behind this aggregate calculation. An aggressive Iraq, contesting the Algiers accord that settled sovereignty over the Shatt El Arab, could determine eighty-three months of war with Iran, and the dispute, settled in August 1990 (Huth, 1996), saw the rivals fight more than half of the time. By contrast, Iran's challenge of the 1937 Accord on the boundaries of the very same Shatt El Arab determined seven brief militarized disputes between 1959 and the signing of the Algiers Accord in 1975, determining a war frequency for the dispute of 20.4%.[20]

The defender's equilibrium response to challenge balances out costs and benefits in expected terms, but this says little about the possible paths that a crisis could take as our rivals implement their equilibrium strategies. The game begins when the challenger expresses dissatisfaction with the *status quo*. This is the point at which he picks a value for s. The dispute can then be in the public eye for a while before a full-fledged challenge, involving a threat of escalation to war, is actually issued. For example, Ecuador's contest of the 1942 Rio Protocol was a strong populist theme domestically before explicit border challenges led to armed encounters in 1953 and 1954. Armed encounters can then be followed by long periods of relative calm before the challenger challenges again. From 1960 to the border incidents of 1978, Ecuador refrained from any overt militarized challenge of the Rio Protocol (Krieg, 1986:224). Nevertheless, the issue remained explicitly unsettled despite long periods of *détente* between Ecuador and Peru.

In our model, the rivals eventually switch roles, and this marks the beginning of a new game. Parameters and strategy determine the possible paths that the rivals could follow as the dispute evolves from the challenger's expressed dissatisfaction with the *status quo* to the defender's submission. Each of the possible paths yields a utility for each rival that measures the discounted sum of benefits and costs received from the start of the game until the rivals switch roles at SB. To illustrate the possible developments that can emerge as a result of equilibrium play, we again set $\omega = 0.99$, and consider the following set of parameter values: $a = 1$, $c =$ 1.5, $b = 2$ and d $= 2$. Moreover, we assume that the challenger chooses $s = 0.3$ and $t = 0.8$. A simulation of the dispute using the strategic probabilities reported in Table 4 is revealing of the many paths that could determine the development of this dispute.[21] The following five paths, described by the sequence of payoff states visited, are representative of the possible payoff outcomes to the rivals:

• $\sigma_1 = \{\text{SB}\}$: The defender immediately submits after a challenge. This happens with probability $P(\sigma_1) = s \times p_1 =$ 0.07 given our parameter values. The utility of this path is 0 for the defender and 0.033 for the challenger. The challenger's utility corresponds to his expectation at SB when he becomes the new defender, and it assumes that the new challenger will, in turn, challenge with probability $s = 0.3$.[22]

• $\sigma_2 = \{\text{SQ, SQ, WR, SQ,SQ,SQ, SB}\}$: The challenger decides to wait for two turns before challenging. The defender therefore enjoys the benefits of possession for the first two turns. The challenger then challenges, and the conflict escalates to war, but then the challenger remains at the *status quo*, waiting three turns before challenging again. The defender submits in the face of this second challenge. The probability of this particular path is $P(\sigma_2) =$ 0.00044, and both challenger and defender receive positive utility from it. Challenger, who has enjoyed the benefits of possession for five turns, receives 0.016 while the defender, who does end up winning the prize having fought one turn of war, receives 0.029.[23]

• $\sigma_3 = \{\text{WR,WR,WR,SB}\}$: The rivals that follow this path immediately escalate the conflict to war, and fight for three periods before the defender submits. If this path obtains, both rivals fight too much relative to the value of switching roles. Their discounted utilities, at -0.013 for the challenger and -0.059 for the defender, are both negative. This is what we refer to as a ruinous outcome for both parties. The probability of this path $P(\sigma_3) =$ 0.003. But the outcome could be ruinous for one and not the other if the rivals travel paths such as σ_4 or σ_5.

• $\sigma_4 = \{\text{SQ}, \text{BD}, \text{WR}, \text{SQ}, \text{BD}, \text{BD}, \text{WR}, \text{BD,SB}\}$: In this path, the challenger backs down in the face of a resisting defender more often than he escalates to war. As a result, the defender enjoys the benefit of a rival backdown more often than she incurs the costs of war. Her payoff is positive at 0.058 while the challenger accumulates audience and war costs in excess of the expected value of the prize at SB. His payoff is -0.037, and this path occurs with probability $P(\sigma_4) =$ 4.81 $\times$ 10^{-6}.

• $\sigma_5 = \{\text{WR,SB}\}$: It is now the defender who experiences a negative payoff outcome. She fights but does not get a chance to enjoy possession of the contested asset before submitting. Her payoff is -0.02 while the challenger's payoff is 0.017. The challenger gets possession fast and without having to fight very much. This path occurs with probability $P(\sigma_5) =$ 0.025.

The probability of any individual path is typically small, but there may be many ways to fight more than the contested asset is worth, or to collect enough rent to cover any outbreak of war. And this will also depend on parameter values and challenger strategy. If each of the catastrophic paths can only happen with minuscule probability, the probability of catastrophic war could remain small even if many paths are associated to this dire outcome. But this need not be the case. In what follows we estimate the likelihood of various payoff outcomes as parameter values and challenger strategy vary.

The Likelihood of Catastrophic War and Other Payoff Outcomes

Strategy is an important determinant of the outcome for our rivals. The challenger can moderate or enhance the intensity of challenge by choice of probabilities s and t. From the perspective of the standard expected utility calculation, the challenger is indifferent between the various degrees of aggression that he can impose on the defender. But the possible outcomes highlight differences in the risks that are taken. Indeed, the challengers' strategic choice, given parameter values, will determine the likelihood of catastrophic war for each party as well as the likelihood that both parties will switch roles having survived the crisis with strictly positive payoffs. These are the downside and upside risks associated with the choices of s and t.

To capture the impact of strategy, we estimated the likelihood of various payoff outcomes using the Monte Carlo method. This is required because the probabilities that we are interested in cannot be calculated explicitly. We estimated the likelihood of various payoff events including that of path $\sigma_1 = \{\text{SB}\}$, in which the challenger immediately challenges the *status quo* and the defender submits. $P(\sigma_1)$ can be calculated explicitly, but by estimating the likelihood of σ_1 separately, we were able to compare one of our probability estimates to a calculated true value. In all cases, our 95% confidence interval for the likelihood of $\sigma_1$contained the true value, and this bolstered our confidence in the interval estimates of the likelihoods of other payoff events of interest.[24] Table 5 below provides some data. In all cases ω is set at 0.99.

<< Table 5 about here >>

An aggressive challenger precipitates ruinous warfare for both sides. This is clear from a comparison of Cases 1 and 2. Increased s and t increases the likelihood that both challenger and defender will fight so much that they will end up with strictly negative payoffs. The challenger fights more than the prize is worth with 50% likelihood if he chooses to challenge the *status quo* with certainty at all relevant opportunities. By contrast, if he adopts restraint and challenges with probabilities $s = t = 0.4$, as in Case 1, he can expect a negative outcome with only 44.4% likelihood.[25] Aggressive action increases the downside risk for the challenger. However, it also dramatically increases the probability that the defender will submit right away ($P(\sigma_1) = .5$ when $s = t = 1$ up from $P(\sigma_1) = .116$ when $s = t = 0.4$).

The increase in the likelihood of a negative outcome for the challenger is part of the defender's deterrent threat. But this comes at a high cost for the defender. If the defender challenges intermittently choosing $s = t = 0.4$ as in Case 1, the defender enjoys a positive outcome with 83.4% likelihood, and with 71.8% likelihood she enjoys a strictly positive utility when the rivals switch roles (since she submits right away with 11.6% probability, and this leaves her with 0 utility). By contrast, countering an aggressive challenger, as in Case 2, diminishes the defender's prospects. Now she receives strictly a positive payoff with 25.1% likelihood only because she submits immediately half of the time. She also faces a negative outcome with higher likelihood when the challenger is aggressive. This is because the challenger prevents her from enjoying possession of the contested asset by always challenging with certainty. As a result, the defender only receives positive payoff from the challenger's occasional backdown. Increased war costs, as expected, increase the likelihood of ruinous warfare for both sides, as illustrated in Cases 3 and 4. But a comparison of the four Cases of Table 5 underscores the importance of strategy. Getting more aggressive increases the challenger's downside risk more than a doubling of his war costs.

In an essay on the comparative merits of theory and case studies, Robert Jervis writes "The choice between the deductive approach and one that builds on case studies involves a tradeoff between rigor and richness. A deductive theory must miss many facets of any individual case." (Jervis, 1989:184). Our theoretical approach to asymmetric deterrence in time is no exception. Nevertheless, our model predicts a wide range of outcomes that are observed in real world settings. Iraq's aggressive posture in the dispute over the Shatt Al Arab led to a war whose direct and indirect costs rose to " the astronomical figure of $1,190 billion" (Hiro, 1991:1). Was sovereignty over the Shatt El Arab waterway worth this much? For sure, wider issues of Arab nationalism and religious fundamentalism added much fuel to a raging fire. But surely war costs accumulated beyond the value of the contested territory. By contrast, the dispute between Mali and Burkino Faso over 500 square miles of territory, including the mineral rich Agacher strip, lasted for seventeen years but escalated to war on only two brief occasions. Throughout the period, Burkino Faso held mineral rich territory of value. And Mali's willingness to challenge and escalate the conflict to war eventually led to an even distribution of the contested territory in 1987 (Huth, 1996:220). It is probable that neither side fought more than the contested asset was worth if exploited, and the defender (Burkino Faso) was able to keep possession for a very long while. Then again, many disputes do not escalate to war at all. Our defender's approach to deterrence can lead to all of these outcomes. General deterrence can succeed given our defender's strategy. But if general deterrence fails, so can immediate deterrence, and the consequences can be ruinous.

CONCLUSION

Our analysis illustrates that unbundling the conflict outcome changes the nature of the deterrent threats that fully informed rivals wield against each other. The defender's strategy in equilibrium makes the challenger indifferent between challenging the *status quo*, with the understanding that war is a possible outcome, or not. General deterrence success is therefore possible. But if the challenger chooses to challenge, then the defender threatens immediate deterrence failure with some probability. War, more or less protracted, is threatened probabilistically by our fully informed rivals. Their strategic behavior can therefore lead to rivalries that "only periodically escalate to the level of militarized conflict and can persist in the absence of such conflicts for a significant period of time," (Goertz and Diehl, 1993:156). In this, our model is suggestive of a strategic foundation to enduring rivalries and predicts some of the observed real world outcomes. More importantly, however, our analysis points to the limitations of a one-shot game analysis of deterrence. If the challenger can challenge the *status quo* repeatedly, immediate deterrence can fail under the very same circumstances that ensured its success in the one-shot framework. Fully informed rivals fight costly wars because the defender will not hand over the contested asset on demand by submitting, but will threaten instead repeated escalation to war in case of challenge. And these wars can last long enough to impose devastating costs.

APPENDIX

The Markov strategies described in the text define a transition matrix T (for each case of c vs a) between the four payoff states $\{$SQ, SB, BD, WR$\}$.[26] If $a \leq c$:

$$T = \begin{pmatrix} 1-s & s(1-p) & sp(1-q) & spq \\ 0 & 1 & 0 & 0 \\ 1-t & t(1-p) & tp(1-q) & tpq \\ 0 & 1-r & r(1-q) & rq \end{pmatrix} \tag{A1}$$

and similarly for $a \geq c$ by exchanging the last two rows of T. For player i (D or C), expectations at these four states satisfy:

$$E_i = U_i + \omega T E_i \qquad \text{or} \qquad [I - \omega T]E_i = U_i \tag{A2}$$

where U_i is the vector of i's payoffs.[27]

Lemma 1: Expectations for the challenger (C) and the defender (D) corresponding to the strategies given by formulae (3), (4), and (5) in the text are:

$$E_C = \begin{pmatrix} 0 \\ \frac{1-\omega}{1-(1-s)\omega} \\ -(1-\omega)a \\ -(1-\omega)a \end{pmatrix} \text{ and } \quad E_D = \begin{pmatrix} \frac{1-\omega}{1-(1-s)\omega} \\ 0 \\ \frac{(1-\omega)\left(\omega(1-t)+b\left(1-\omega(1-s)\right)\right)}{1-(1-s)\omega} \\ -(1-w)d \end{pmatrix} \text{ if } a \leq c\text{; and}$$

$$E_C = \begin{pmatrix} 0 \\ \frac{1-\omega}{1-(1-s)\omega} \\ -(1-\omega)c \\ -(1-\omega)c \end{pmatrix} \text{ and } \quad E_D = \begin{pmatrix} \frac{1-\omega}{1-(1-s)\omega} \\ 0 \\ (1-w)b \\ \frac{(1-\omega)\left(\omega(1-t)-d\left(1-\omega(1-s)\right)\right)}{1-(1-s)\omega} \end{pmatrix} \text{ if } a \geq c.$$

Proof: One verifies (A2) by writing T and E_i as given in (A1) and the lemma. For instance, in the case $a \leq c$, the first row of (A2) for C reads in dot product form:

$$< 1-\omega(1-s),\ -\omega s(1-p),\ -\omega sp(1-q),\ -\omega spq > . < 0, \tfrac{1-\omega}{1-(1-s)\omega},\ -(1-\omega)a,\ -(1-\omega)a >$$

$$= -\omega s(1-p)\tfrac{1-\omega}{1-(1-s)\omega} + \omega sp(1-\omega)a = \omega s(1-w)\left(ap - \tfrac{1-p}{1-(1-s)\omega}\right) = 0$$

since $ap - \frac{1-p}{1-(1-s)\omega} = 0$ according to $p = p_1$ in section (5a). All other cases are similar.[28] *Q.E.D.*

Theorem 1: Formulae (3), (4), and (5) in the text provide a MPE with $p = p_1,\ q = q_1,$ and $r = r_1$

if $a \le c \le a + \frac{\omega}{1-\omega(1-s)}$ (A3)

and with $p = p_2,\ q = q_2$, and $r = r_2$

if $c \le a \le c + \frac{\omega}{1-\omega(1-s)}$ and $\frac{\omega(1-t)}{1-\omega(1-s)} \le d$ (A4)

Proof: All formulae for $p,\ q,$ and r must provide true probabilities. This is obvious for p_1 and p_2. To ensure, for instance, that r_1 is a probability one simply solves the inequalities $0 \le r_1 \le 1$ which result in the right-hand side of (A3). All other conditions are obtained similarly. In addition, one must ensure that Challenge is indeed best at WR1 in the case $a \le c$ and at BD1 in the case $a \ge c$. In the first case, for instance, since Challenger's expectation for waiting is $E_C(1) = 0$ the expectation for challenging must be non-negative, or:

$$(1 - r_1)E_C(2) + r_1\big((1 - q_1)E_C(3) + q_1E_C(4)\big) \ge 0 \quad \text{(A5)}$$

This reduces to: $(1 - r_1)\frac{1-\omega}{1-\omega(1-s)} - r_1(1 - \omega)a \ge 0,$ or $a \le c$, after replacing r_1 as in (5c). The other case is similar. Moreover, since $E_C(3) = E_C(4)$ in both cases of a vs c, Challenger's use of probability q is optimal. Finally, Defender's use of probabilities p and r is optimal provided that:

$$E_D(2) = (1 - q)E_D(3) + qE_D(4) \quad \text{(A6)}$$

In the case $a \le c$, for instance, this reads:

$$0 = (1 - q_1)\frac{(1-\omega)\Big(\omega(1-t)+b\big(1-\omega(1-s)\big)\Big)}{1-(1-s)\omega} - q_1(1 - w)d$$

which yields formula (4a) for q_1. The other case is similar.[29] *Q.E.D.*

REFERENCES

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

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

Bueno de Mesquita, Bruce and David Lalman (1992) *War and Reason*. New Haven: Yale University Press.

Downs, George and David Rocke (1994) "Conflict, Agency, and Gambling for Resurrection: The Principle-Agent Problem Goes to War," *American Journal of Political Science*, 38:362-80.

Fearon, James (1995) "Rationalist Explanations for War," *International Organization* 49: 379-414.

Fearon, James (2002) "Why Do Some Civil Wars Last So Long?" Manuscript, Department of Political Science, Stanford University.

Filson, Darren and Suzanne Werner (2002) "A bargaining Model of War and Peace: Anticipating the Onset, Duration and Outcome of War," *American Journal of Political Science* 46:819-38

Flynn, Todd (1994) "Limited War and Regional Conflict Escalation," Ph.D. dissertation, Department of Political Science, Stanford University.

Fudenberg, Drew and Jean Tirole (1995) *Game Theory*, The MIT Press, Cambridge, MA.

Garfinkel, Michelle and Stergios Skaperdas (2000) "Conflict without Misperceptions or Incomplete Information: How Future Matters," Journal of Conflict Resolution 44:793-807.

Goertz, Gary and Paul Diehl (1998) "The Political Uncertainty of Interstate Rivalries: A Punctuated Equilibrium Model" in *The Dynamics of Enduring Rivalries*, Paul Diehl Editor. Urbana: University of Illinois Press.

Goertz, Gary and Paul Diehl (1993) "Enduring Rivalries: Theoretical Constructs and Empirical Patterns," *International Studies Quarterly* 37:147-171.

Heldt, Birger (1999) "Domestic politics, Absolute Deprivation and the Use of Armed Force in Interstate Territorial Disputes, 1950-1990," *Journal of Conflict Resolution* 43:451-478.

Hiro, Dilip (1991) *The Longest War, The Iran-Iraq Military Conflict*. New York: Routledge, Clapman and Hall Inc.

Huth, Paul (1996) *Standing Your Ground. Territorial Dispute and International Conflict*. Ann Arbor: University of Michigan Press.

Huth, Paul and Bruce Russett (1993) "General Deterrence Between Enduring Rivals: Testing Three Competing Models," *American Political Science Review* 87:61-73.

Jervis, Robert (1984) *The Illogic of American Nuclear Strategy*. Ithaca and London: Cornell University Press.

Jervis, Robert (1989) "Rational Deterrence: Theory and Practice," *World Politics* 41:183-207.

Johnston, Alastair (1995) *Cultural Realism: Strategic Culture and Grand Strategy in Chinese History*. Princeton, New Jersey: Princeton University Press.

Kocs, Stephen (1995) "Territorial Disputes and Interstate War, 1945-1987," *Journal of Politics* 57:653-73.

Krieg, William (1986) *Ecuadorean-Peruvian Rivalry in the Upper Amazon*. Washington DC. Department of State.

Kydd, Andrew (1997) "Game Theory and the Spiral Model," *World Politics* 49:371-400.

Langlois, Jean-Pierre (1991) "Rational Deterrence and Crisis Stability," *American Journal of Political Science* 35:801-32.

Lemke, Douglas and William Reed (2001) "War and Rivalry Among Great Powers," *American Journal of Political Science* 45:457-469.

Lieberman, Elli (1995) "What Makes Deterrence Work? Lessons from the Egyptian-Israeli Enduring Rivalry," *Security Studies* 4:851-910.

Maoz, Zeev and Ben Mor (2002) *Bound by Struggle: The Strategic Evolution of Enduring International Rivalries*. Ann Arbor: Michigan University Press.

Maoz, Zeev and Ben Mor (1996) "Enduring Rivalries: The Early Years," *International Political Science Review* 17:141-160.

Mooney, Christopher (1997) *Monte Carlo Simulation*. Quantitative Applications in the Social Sciences Series, Sage University Paper 116. Thousand Oaks, London, New Delhi: Sage Publications.

Morgan, Clifton (1977) *Deterrence, A Conceptual Analysis*. Beverly Hills, CA: Sage.

Morrow, James (1989) "Capabilities, Uncertainty and Resolve: A Limited Information Model of Crisis Bargaining, " *American Journal of Political Science* 83:941-972.

Morrow, James (1997) "When do Relative Gains Impede Trade?," *Journal of Conflict Resolution* 41:12-37.

Morrow, James (2000) "The On-Going Game Theoretic Revolution," in Manus I. Midlarsky Ed., *Handbook of War Studies II.* Ann Arbor: University of Michigan Press.

O'Neill, Barry (1994) "Game Theory Models of Peace and War," in Robert J. Aumann and Sergui Hart Eds. *Handbook of Game Theory with Economic Applications, Vol. II*. Amsterdam: Elseiver.

O'Neill, Barry (1989) "Game Theory and the Study of Deterrence of War," in Paul C. stern, Robert Axelrod, Robert Jervis and Roy Radner Eds. *Perspectives on Deterrence*. New York: Oxford University Press.

Ordershook, Peter (1989) *Models of Strategic Choice in Politics*. Ann arbor: University of Michigan Press.

Powell, Robert (1987) "Crisis Bargaining, Escalation and MAD," *American Political Science Review* 81:717-35.

Powell, Robert (1990) *Nuclear Deterrence Theory: The Search for Credibility*. New York: Cambridge University Press.

Powell, Robert (1996a) "Uncertainty, Shifting Power, and Appeasement," *American Political Science Review* 90:749-764.

Powell, Robert (1996b) "Stability and the Distribution of Power," *World Politics* 48:239-267.

Powell, Robert (1999) *In the Shadow of Power*. Princeton: Princeton University Press.

Powell, Robert (2003) "Bargaining and Learning while Fighting," Manuscript, department of Political Science, UC Berkeley.

Reiser, Stewart (1994) "The Arab-israeli Wars: A Conflict of Strategic Attrition," in *Prolonged Wars, A Post Nuclear Challenge*. Karl P. Magyar and Constantine P. Danopoulos, Eds., Maxwell Airforce Base. Alabama: Air University Press.

Sartori, Anne (2003) "Enduring Facts About Enduring Rivals," Manuscript, Princeton University.

Schelling, Thomas (1960) *The Strategy of Conflict*. Cambridge MA: Harvard University Press.

Schultz, Kenneth (1999) "Do Democratic Institutions Constrain or Inform," *International Organization* 53:23-66.

Schultz, Kenneth (2001) *Democracy and Coercive Diplomacy*. New York: Cambridge University Press.

Simmons, Beth (1999) "Territorial Disputes and Their Resolution: The Case of Ecuador and Peru,"Peaceworks No. 27, April, United State Institute for Peace, Washington DC.

Slantchev, Branislav (2003) "The power to Hurt: Costly Conflict with Completely Informed States," *American Political Science Review* 97:123-135.

Smith, Alastair (1998a) "Fighting Battles, Winning Wars," *Journal of Conflict Resolution* 42:301-320.

Smith, Alastair (1998b) "International Crises and Domestic Politics," *American Political Science Review* 92:623-638.

Smith, Alastair (1996) "Diversionary Foreign Policy in Democratic Systems," *International Studies Quarterly* 40:133-153.

Smith, Alastair and Alan Stam (2001) "Bargaining Through Conflict," Unpublished manuscript, Department of Political Science, Yale University.

Snyder, Glenn (1964) "The Balance of Power and the Balance of Terror" in Paul Seabury, ed., *The Balance of Power*. San Francisco: Chandler.

Sobol, Ilya (1994) *A Primer for the Monte Carlo Method*. Boca Raton, Ann Arbor, London, Tokyo: CRC Press.

Vasquez, John (1996) "Distinguishing Rivals That Go to War From Those That Do Not: A Quantitative, Comparative Study of the Two Paths to War," *International studies Quarterly* 40:531-558.

Vasquez, John (1995) "Why do Neighbors Fight? Proximity, Interaction of Territoriality," *Journal of Peace Research* 32:277-293.

Wagner, Harrison (1992) "Rationality and Misperception in Deterrence Theory," *Journal of Theoretical Politics* 4:115-141.

Wagner, Harrison (2000) "Bargaining and War," *American Journal of Political Science* 44:469-484.

Weidner, Glenn, R. (1996) "Operation Safe Border: The Ecuador-Peru Crisis," *Joint Forces Quarterly*, Spring.

Zagare, Frank (1990) "Rationality and Deterrence," *World Politics* 42:238-260.

Zagare, Frank (2004) "Reconciling Rationality with Deterrence: A Re-examination of the Logical Foundations of Deterrence Theory," *Journal of Theoretical Politics* 16:107-141.

Zagare, Frank and Marc Kilgour (2000) *Perfect Deterrence*. Cambridge Studies in International Relations. Cambridge: Cambridge University Press.

Zagare, Frank and Mark Kilgour (1998) "Deterrence Theory and the Spiral Model Revisited," *Journal of Theoretical Politics* 10:66-87.

Zagare, Frank and Mark Kilgour (1993) "Asymmetric Deterrence," International Studies Quarterly 37:1-27.

NOTES

[1]War is still a game ending move if the model includes an escalation ladder as in Bueno de Mesquita and Lalman (1992), Kydd (1997) or Zagare and Kilgour (2000). In the asymmetric case fully informed rivals then never escalate to the highest level of conflict as long as the *status quo* is preferred to war. But in a symmetric case where either side can challenge, this is not necessarily true as shown in Bueno de Mesquita and Lalman (1992:72-75)

[2]We do not make any claim that our study is exhaustive. Our aim is simply to point out a range of completely rational interactions that contradicts the conventional wisdom on when deterrence succeeds.

[3]In the full information models developed by Slantchev (2003) and Garfinkel and Skaperdas (2000), the accumulated costs of war can never exceed the value of the prize. Our model therefore illustrates, as does Flynn (1994), the possible catastrophic outcome of repeated challenge and escalation to war by fully informed rivals.

[4]Goertz and Diehl (1998) develop a punctuated equilibrium model of enduring rivalries. Inspired by the biology literature, these authors use the punctuated equilibrium analogy to interpret enduring rivalries as a stable phenomenon in time interrupted by shocks that mark the beginning and end of these relationships. Such a model appeals to evolutionary ideas but does not address the possible strategic roots of rivalry.

[5]See also, for example, Vasquez (1995) or Kocs (1995) on the importance of territory to explain war.

[6]This is true of the models developed by Fearon (1994a, 1994b,1995), Morrow (1989, 1997), Powell (1996a, 1996b, 1999), Schultz (1999, 2001), and Smith (1998b).

[7]The model is an extension of the model in Smith (1998a) in which the likelihood that the battle for a fort will be won is exogenous.

[8]Fully informed rivals have also been shown to choose conflict as a result of commitment problems or indivisibilities. See, for example, Fearon (1995, 2002).

[9]Asymmetric deterrence is understood, following Zagare and Kilgour (1993), as a situation which involves "one decision maker, "Challenger," (who) decides whether to initiate a crisis involving a second decision maker, "Defender." (Zagare and Kilgour, 1993, p.1). The one-shot model we discuss here is identical to the one examined by these authors. Our normalization of payoffs is somewhat different but is designed to facilitate the analysis of our generalization.

[10]To our knowledge there is no universally accepted terminology for our game structure. Although it is *infinitely* iterated, it is not a classical "repeated game" since the structure changes according to prior developments. It is in fact a game on a graph and could also be called a "stochastic game" because the transition between its states defines a Markov chain.

[11]Although general deterrence can succeed in this case, there is no guarantee that the rivals will not fight. Indeed, one could construct equilibria based on reversion to extremal equilibria, (see Slantchev, 2003) in which the parties agree to fight for some time before returning to the *status quo*. However, such equilibria are not directly relevant to our discussion of deterrence since they illustrate how rivals could be deterred from *refusing* to fight by a threat of reversion to their worst outcome.

[12]Our discussion of the cost of limited war illustrates that it is *per period* payoffs that are conceptually important to the rivals in the iterated game of Figure 2. Thus, for example, the challenger considers the *per period* cost of war $(1-\omega)c$ as it relates to $(1-\omega)a$ when deciding on strategy. But comparing these *per period* costs is of course equivalent to comparing parameters c and a directly. In what follows we can therefore refer to the parameter values themselves in discussing strategy, since the *conceptually* relevant *per period* payoffs are, given ω, a fixed fraction of these parameter values.

[13]Details of these arguments are available from the authors upon request.

[14]The upper bounds for s (and c) are given in Theorem 1, formulae (A3) and (A4) in appendix.

[15]When $a \geq c$, probabilities s and t must satisfy: $t \geq 1 - \frac{1}{\omega}\big(1 - \omega(1-s)\big)d$.

[16]Following (A3) and (A4) in appendix, upper bounds for s are calculated using the formula $s \leq \frac{1}{|c-a|} - \frac{1}{\omega} + 1$. Lower bounds for t use the formula of footnote 13.

[17]It might seem surprising that the challenger would use the very same probability q at all three nodes of Figure 2 where the decision to escalate arises. But this is because of the simple structure implicit in that figure: we only distinguish three possible pasts (SQ if challenger just waited, BD if he just backed down, and WR if he just escalated). In order to allow an equilibrium that involves probabilities for the defender to resist (at SQ2, BD2, and WR2), it is necessary to make her indifferent between going to SB by submitting, with a known result (0 forever), and going to each of SQ3, BD3, and WR3. But her expectation at all three nodes is solely determined by her expectations at BD and WR and the probabilities of escalation used by the challenger. So, these probabilities must be a same q in order to make the defender indifferent with submitting.

[18]The rivals expectations in each of the four states SQ, SB, BD and WR are described in Lemma 1 in Appendix.

[19]Consider the case where $c = 1.5$. Comparing the data in the first column of Table 4 note that, when $a = 1$, the defender submits at BD2 with probability $(1-p) = 0.33$, but when $a = 4$, the defender submits at BD2 with probability $(1-r) = 0.85$.

[20]Huth (1996, p.211) reports that Iran and Iraq fought for a total of 41 months between November 1959 and June 1975 or 20.4% of the time.

[21]Technically, an infinite number of paths are possible but, of course, the probability that the crisis will develop along a particular path decreases with its length.

[22]Recall that the challenger's utility at SB is $\frac{1-\omega}{1-\omega+\omega s}$, which in this case is equal to $\frac{0.01}{0.01+0.99\times 0.3} = 0.033$.

[23]Utilities are computed as follows: For defender 0.029 $= .01 + ,99 \times .01 - .99^2 \times .02 + .99^3 \times .01 + .99^4 \times .01 + .99^5 \times .0 + 0$ since once the defender submits, she

becomes the challenger and has 0 expected utility at SB. For the challenger $0.016 = 0 + 0 - .99^2 \times .015 + 0 + 0 + 0 + .99^6 \times \frac{.01}{.01+.99\times.3}$ since when the challenger becomes the new defender at SB he expects a payoff of $\frac{1-\omega}{1-\omega+\omega s}$.

[24]Having consulted Mooney (1997) and Sobol (1994), we implemented the Monte Carlo method as follows: for each set of parameter values, we generated 20,000 paths by simulating the rivals' equilibrium play. We calculated the players' *ex ante* utilities for each path. We first sorted the paths into four groups: paths for which both parties receive strictly negative payoffs; paths for which the players received positive payoffs; paths for which one or the other receives a strictly negative payoff. Frequencies for each of the four events were calculated. We also further subdivided the group of paths for which the rivals receive positive utilities by separating path σ_1, for which the defender receives 0 utility, from all other paths. In practice both defender and challenger receive *strictly* positive payoffs for these remaining paths.(In theory it is possible that one of the rivals could receive exactly 0 by traveling one of these paths. This would require that discounted costs exactly compensated for discounted benefits, an event which is highly unlikely). This process was repeated 1000 times, generating a distribution for the frequencies of each event. In all cases, with 95% confidence, we were not able to reject the null hypothesis of a normal distribution (Jarque Bera statistic $<$ 5.99) . We were therefore able to construct 95% confidence interval estimates of the likelihoods of each of the six events by adding and subtracting 1.96 standard deviations to each of the means.

[25]44.4% adds the mean likelihoods in the second column of the Case 1 matrix. Since we cannot reject the hypothesis that the estimated likelihoods are normally distributed, we cannot reject the hypotheses that their sum is either. We can therefore calculate confidence intervals for these sums. When $s = 0.4$, $t = 0.4$, the 95% confidence interval for the likelihood that the challenger's utility is strictly negative is [.436, .452]. When $s = 1$, $t = 1$, the 95% confidence interval for the likelihood that the challenger's utility is strictly negative is [.491, .509]. The two statistics are therefore significantly different.

[26]Note that SB is treated here as an "absorbing state" that ends the game although it is in fact the beginning of a new game with the players changing hats. This simplifies the mathematics since the expected value for one player at SB is simply the expected value of the other at SQ given the assumed symmetry in the two games.

[27]$U_C = (1-\omega)\begin{pmatrix} 0 \\ \frac{1-\omega}{1-(1-s)\omega} \\ -a \\ -c \end{pmatrix}$ and $U_D = (1-\omega)\begin{pmatrix} 1 \\ 0 \\ b \\ -d \end{pmatrix}$

[28]Details of all calculations are available from the authors upon request.

[29]Inequalities (A3) and (A4) are constraints on the relative magnitudes of s and t.

FIGURES

Figure 1: The One Shot Asymmetric Deterrence Game

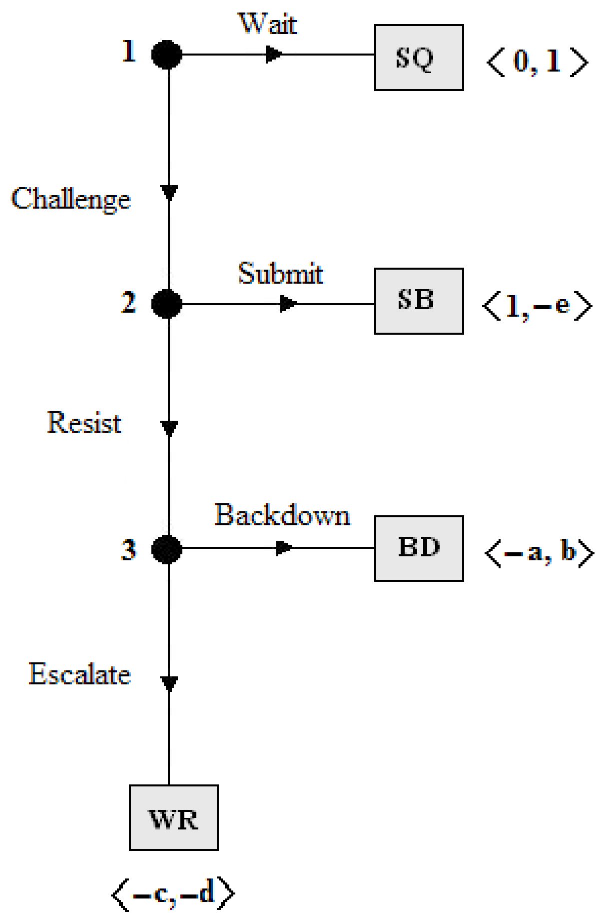
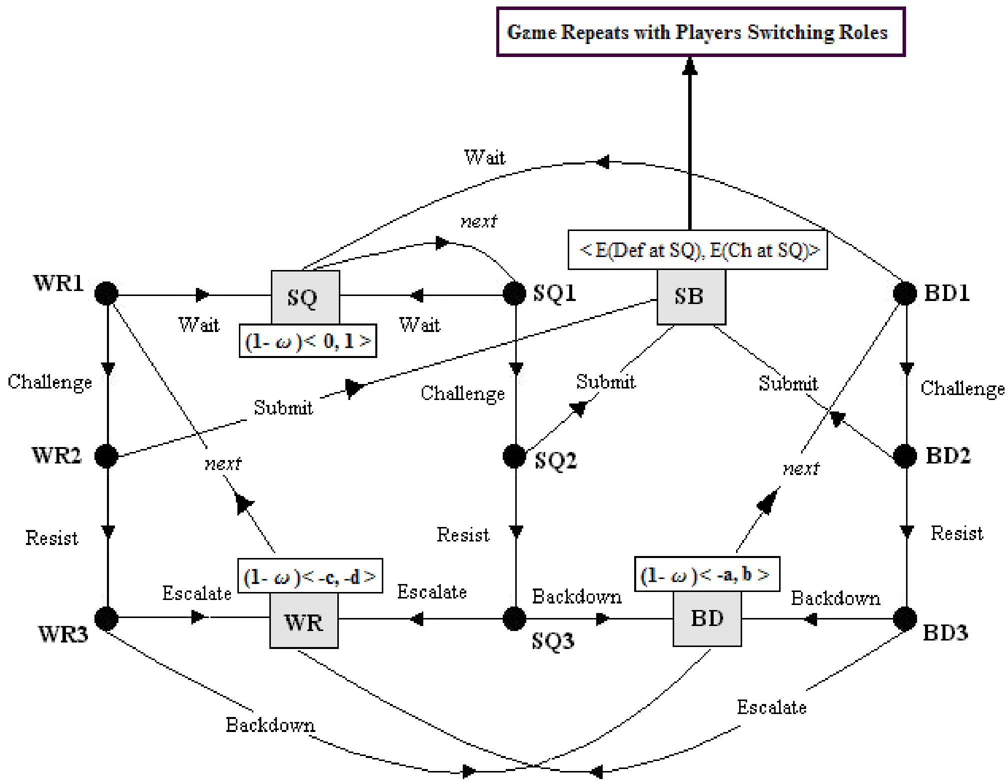


Figure 2: The Discounted Iterated Game with Four States



TABLES

Table 1: Probabilities of Challenge at SQ1, BD1 and WR1

		$c = 1.5$		$c = 3.0$
$a = 1$	**At SQ1**	$0 \leq s \leq 1.00$	**At SQ1**	$0 \leq s \leq 0.49$
	At BD1	$0 \leq t \leq 1.00$	**At BD1**	$0 \leq t \leq 1.00$
	At WR1	1	**At WR1**	1
$a = 4$	**At SQ1**	$0 \leq s \leq 0.39$	**At SQ1**	$0 \leq s \leq 0.99$
	At BD1	1	**At BD1**	1
	At WR1	$t \geq 0.98 - 2s$	**At WR1**	$t \geq 0.98 - 2s$

Table 2: The challenger's Probability of Escalation at SQ3,BD3 and WR3

	$\boldsymbol{s = 0.2}$	$\boldsymbol{s = 1}$
$\boldsymbol{t = 0.4}$	$q = 0.71$	$q = 0.56$
$\boldsymbol{t = 0.8}$	$q = 0.60$	$q = 0.52$

Table 3: The Defender's Decision to Resist in Case of Challenge

$a = 1,\ c = 1.5$	
If challenger chooses $s = 0.2$	**If challenger chooses $s = 1$**
Defender resists:	**Defender resists:**
At **SQ2** with probability 0.83	At **SQ2** with probability 0.5
At **BD2** with probability 0.83	At **BD2** with probability 0.5
At **WR2** with probability 0.74	At **WR2** with probability 0.24
$a = 4,\ c = 1.5$	
If challenger chooses $s = 0.2$	**If challenger chooses $s = 0.39$**
Defender resists:	**Defender resists:**
At **SQ2** with probability 0.76	At **SQ2** with probability 0.63
At **BD2** with probability 0.36	At **BD2** with probability 0
At **WR2** with probability 0.76	At **WR2** with probability 0.63

Table 4: Length of the Crisis and Frequency of War Given a Challenge

$s = 0.3$, $t = 0.8$					
Case 1	$c = 1.5$		**Case 2**	$c = 3.0$	
$a = 1$	At SQ2, BD2 $p = 0.77$ At WR2 $r = 0.64$		$a = 1$	At SQ2, BD2 $p = 0.77$, At WR2 $r = 0.29$	
	Challenger escalates with probability $q = 0.57$			Challenger escalates with probability $q = 0.57$	
	Expected Dispute Length	6.59		**Expected Dispute Length**	5.28
	Expected Turns of War	1.44		**Expected Turns of War**	0.86
	War frequency	21.9%		**War frequency**	16.3%
Case 3	$c = 1.5$		**Case 4**	$c = 3.0$	
$a = 4$	At WR2, SQ2 $p = 0.68$ At BD2 $r = 0.15$		$a = 4$	At WR2, SQ2 $p = 0.52$, At BD2 $r = 0.36$	
	Challenger escalates with probability $q = 0.60$			Challenger escalates with probability $q = 0.60$	
	Expected Dispute Length	5.14		**Expected Dispute Length**	4.67
	Expected Turns of War	0.77		**Expected Turns of War**	0.57
	War frequency	15.0%		**War frequency**	12.2%

Table 5: Likelihood of ruinous warfare and Other Payoff Outcomes

Impact of Strategy Given $\boldsymbol{a = 1,\ c = 1.5,\ d = 2,\ b = 2}$

Case 1 $s = 0.4, t = 0.4$

	Challenger Utility	
Defender Utility	≥ 0	< 0
≥ 0	[.495, .509] Mean .502	[.321, .334] Mean .328
< 0	[.050, .056] Mean .053	**[.112, .120]** **Mean .116**
$P(\sigma_1) \in$ [.112, .120] True value .116		

Case 2 $s = 1, t = 1$

	Challenger Utility	
Defender Utility	≥ 0	< 0
≥ 0	[.493, .507] Mean .50	[.245,.257] Mean .251
< 0	No cases Mean 0	**[.242, .256]** **Mean .249**
$P(\sigma_1) \in$ [.493, .507] True value .50		

Impact of War Costs Given Strategy $\boldsymbol{s = 0.4,\ t = 0.4}$

Case 3 a $= 1,\ \boldsymbol{c = 3},\ d = 2,\ b = 2$

	Challenger Utility	
Defender Utility	≥ 0	< 0
≥ 0	[.519, .533] Mean .526	[.302, .314] Mean .308
< 0	No cases Mean 0	**[.161, .171]** **Mean .166**
$P(\sigma_1) \in$ [.112, .120] True value .116		

Case 4 a $= 1,\ c = 1.5,\ \boldsymbol{d = 4},\ b = 2$

	Challenger Utility	
Defender Utility	≥ 0	< 0
≥ 0	[.526, .540] Mean .523	[.203, .215] Mean .209
< 0	[.064, .070] Mean .067	**[.185, .197]** **Mean .191**
$P(\sigma_1) \in$ [.111, .121] True value .116		