International Studies Quarterly

International Studies Quarterly (2011) **55**, 1–24

The Dark Side of the Future: An Experimental Test of Commitment Problems in Bargaining[1]

Dustin H. Tingley

Harvard University

While most existing theoretical and experimental literatures focus on how a high probability of repeated play can lead to more socially efficient outcomes (for instance, using the result that cooperation is possible in a repeated prisoner's dilemma), this paper focuses on the detrimental effects of repeated play—the ''dark side of the future.'' I study a resource division model with repeated interaction and changes in bargaining strength. The model predicts a negative relationship between the likelihood of repeated interaction and social efficiency. This is because the longer shadow of the future exacerbates commitment problems created by changes in bargaining strength. I test and find support for the model using incentivized laboratory experiments. Increases in the likelihood of repeated play lead to more socially inefficient outcomes in the laboratory.

Two major themes in the existing theoretical and empirical literatures on resource division problems are the likelihood of continued interaction (''the shadow of the future'') and changes in the relative bargaining strength of the players. I ask whether commitment problems that stem from changes in bargaining strength (Powell 2006) are exacerbated by a longer shadow of the future. I study this question by analyzing a repeated resource division problem with changes in bargaining strength, where I experimentally manipulate the likelihood of continued interaction. While international relations scholars have frequently been interested in repeated interaction (for example, Axelrod 1984; Axelrod and Keohane 1985; Oye 1985) and changes in bargaining strength (for example, Kim and Morrow 1992; Powell 2006), there is little work that examines them in conjunction. Using incentivized laboratory experiments, I find support for game-theoretic predictions that increasing the shadow of the future leads to costly bargaining breakdowns. That is, the commitment problem highlighted by others as a source of conflict can be exacerbated by a longer shadow of the future.

Changes in bargaining strength can create an incentive to reject offers in favor of a costly alternative. Previous work in IR identifies these incentives and develops the concept of preventive war, whereby a state in decline chooses to launch an

[1] I gratefully acknowledge financial support from Princeton's Bobst Center. I am indebted to Scott Ashworth, Sarah Bermeo, Thomas Christensen, Hein Goemans, Joanne Gowa, Alice Hsiaw, Kosuke Imai, Robert Keohane, Rose McDermott, Adam Meirowitz, Helen Milner, Andrew Moravcsik, Kristopher Ramsay, Jacob Shapiro, and seminar audiences at University of California at San Diego, William and Mary, Harvard University, Texas A&M, and the University of Rochester for discussions, criticisms, and encouragement. Thanks to Chris Crabbe for help with Multistage programming. Replication data available at http://hdl.handle.net/1902.1/14820.

doi: 10.1111/j.1468-2478.2011.00654.x

attack while it retains a bargaining advantage (for example, Levy 1987; Kim and Morrow 1992; Schweller 1992; Fearon 1995; Powell 1996; Powell 1999; Powell 2006).[2] Fearon's (1995) seminal work on rational choice explanations for war, whose model I draw on, offers future changes in bargaining strength as an explanation for preventive war. I study the relationship between repeated interaction and changes in bargaining strength by modeling equilibrium behavior in a repeated resource division problem with an exogenous change in bargaining strength. The actor who anticipates an increase in bargaining strength faces a commitment problem, in that he cannot credibly promise that he will not take advantage of his future strength. Furthermore, increases in the likelihood of continued interaction make the socially inefficient outcome of war more likely. An actor that could be disadvantaged by a change in bargaining strength in the future has a greater incentive to choose a socially inefficient strategy, the more he values the future or believes the disadvantageous future period of negotiation will be reached. This result contrasts with the focus of most substantive and experimental work that sees a positive relationship between the shadow of the future and socially efficient outcomes. Hence, the spirit of my work is similar to observations in Skaperdas and Syropoulos (1996), Fearon (1998), Garfinkel and Skaperdas (2000), Skaperdas (2002), Barkin (2004), Streich and Levy (2007), and Toft (2007).

The negative relationship between social efficiency and repeated interaction that I study contrasts with a more conventional account that sees a positive relationship. Work in strategic decision-making experiments (for example, Andreoni and Miller 1993; Dal Bo 2005) and IR (for example, Axelrod 1984; Axelrod and Keohane 1985; Oye 1985; Snidal 1985; Keohane 1986) often stress a positive relationship between the likelihood of repeated interaction and socially efficient outcomes. A longer ''shadow of the future'' leads to more cooperation. This has led theorists to suggest ways to increase the shadow of the future, such as decomposing bargaining over time or linking issue areas together (Oye 1985:17). Axelrod (1984) argues that the prolonged interaction between trench warfare units in World War I led to more cooperative behavior. And Van Evera (1985) argues that short-term concerns about shifting power lead to a breakdown in cooperation in Europe in 1914, though long-term considerations also weighed on decision making.[3] However, these literatures focus on incentive structures akin to the prisoner's dilemma. Not all interactions in IR can be characterized as a PD game, such as those that are zero-sum in nature, and conclusions based on this game do not necessarily port to other situations.[4] I study a model where increases in the likelihood of future interaction lead to less efficient outcomes: ''the dark side of the future''. When an opponent will be in a better bargaining position in the future, *and* it is likely a decision maker will have to bargain with that opponent in the future, it can be rational to initiate a conflict in the current period when the decision maker's bargaining strength is more advantageous.

[2] A related literature in IR is power transition theory (for example, Tammen et al. 2000; Lemke 2002; Rapkin and Thompson 2003). There is theoretical (Gartzke 1999) and empirical debate (Lemke 2003) about the explanatory power of preventive war.

[3] ''Europe's misperceptions created a strategic setting that was inimical to international cooperation. The shadow of the future was short because many Europeans expected that the final control of Europe would soon be decided, either by war or by peaceful shifts in the balance of power. This prospect reflected the effects of the cult of the offensive, the bandwagon notions, and the tendency to exaggerate the hostility of others, which together suggested that decisive wars or arms races lay in the near future'' (Van Evera 1985:82; see also 1985:108–110). Interestingly, Van Evera also emphasizes more distant problems involving the long-term control of Europe and hence a longer shadow of the future (see also Fischer 1961). Van Evera's focus on other causal mechanisms, such as private information, are beyond the scope of this paper.

[4] Streich and Levy (2007:199) also note this contrast in reference to a bargaining environment similar to what I study: ''(t)he greater the weight a declining actor gives to the future, the greater its incentives for war now. To take another example, the ability to sustain cooperation in an iterated Prisoner's Dilemma game depends on actors having a sufficiently large discount factor or shadow of the future.''

Existing empirical work in IR that tests the influence of changes in bargaining strength uses large-N statistical analysis of country-level dyadic conflicts (for example, Lemke 2002), field data from subnational conflicts (Fearon and Laitin 2008), or detailed case studies (for example, Levy 1987). These studies do not consider expectations about the probability of repeated interaction in the future. Ripsman and Levy (2007) note the role of the shadow of the future but find it hard to study, and their case study does not engage the theoretical connection between changes in bargaining strength and the shadow of the future. Toft (2007) describes how asymmetries in time horizons could lead to civil wars. In his discussion of why democracies might be less likely to launch preventive wars, Levy (2008:9) also notes the possible connection between the propensity for preventive war and the time horizons of democratic versus autocratic leaders but does not develop this point in the historical analysis.[5] One might also consider differences in the incentives of term-limited versus non-term-limited democratic leaders, or democratic leaders in their first (where they might have longer shadow of the future) versus the final term (where they have shorter time horizons) (Hess and Orphanides 1995). While there appears to be interest among IR scholars in how the shadow of the future and incentives for preventive war interrelate, there is little systematic empirical work. A number of scholars note that this is due to difficulties in measuring the shadow of the future (Fearon 1998; Ripsman and Levy 2007; Levy 2008). The inability of existing observational data to test causal arguments is an important reason to use experiments (Morton and Williams 2010), and this paper helps fill this lacuna using incentivized laboratory experiments.

There is no work that tests the predictions of the Fearon model about the influence of the ''shadow of the future''. I provide a test using incentivized laboratory experiments and find substantial support for the model's comparative static predictions. When there are future changes in relative bargaining strength, the probability of ongoing interaction has an important influence on subject behavior. The higher this probability of continued interaction is, the more likely it is that the commitment problem created by changes in bargaining strength leads to inefficient outcomes. This is particularly the case after subjects have more experience with the game. Along with the results from a separate working paper pursued independently and using a different model (McBride and Skaperdas 2009), this is (to the author's knowledge) the first experimental demonstration of a negative relationship between the shadow of the future and social efficiency.

The structure of the paper is as follows. Existing Experimental Work relates the paper to the existing experimental literature on (i) repeated play and (ii) bargaining problems. Model presents a simple model, offered by Fearon (1995), as an explanation for ''preventive war'' and calculates simple comparative statics on the magnitude of power shifts and the likelihood of ongoing interaction. Experimental Design introduces a research design that allows me to study these theoretical predictions using incentivized laboratory experiments. Results presents results based on hypotheses from the game-theoretic model and Conclusion concludes.

Existing Experimental Work

Experiments on Repeated Interaction

A strong tradition in the experimental literature is the study of how repeated interaction, and changes in the likelihood of continued interaction, influences

[5] Olson (2000:42) raises a similar point about the differences in time horizons between autocratic and democratic leaders.

behavior. Early experiments on repeated prisoner's dilemma-type games found relatively mixed results about the ''shadow of the future.'' Roth and Murnighan (1978) find some support for the idea that repetition leads to more cooperative behavior, but small changes to the incentives of the PD game produce conflicting results. Palfrey and Rosenthal (1994) study public good contributions but find relatively small differences between single shot and repeated play versions of the experiment.[6] On the other hand, Duffy and Ochs (2009) find more support for a connection between repeated play and cooperation, but note the paucity of experiments on infinitely repeated games. In a recent paper, Dal Bo (2005) compares repeated games of fixed, known length to repeated games of uncertain length but with expectation equal to the fixed round version. He finds that the effect of repeated play is indeed due to the probability of continued interaction (''shadow of the future'') and not simply the greater number of rounds to be played. In contrast, I use the discounted infinitely repeated play framework but focus on how repetition can lead to more socially inefficient outcomes and use a resource division game instead of a prisoner's dilemma game.

Experiments on Bargaining Strength

Experiments on bargaining strength typically do not allow bargaining strength to change permanently. In a classic study, Binmore, Morgan, Shaked, and Sutton (1991) investigate behavior in a two-person divide-the-pie game with alternating offers where, if bargaining ends, each player receives a ''breakdown'' payoff that can differ across players. These breakdown payoffs determine the relative bargaining strength of the individual. In contrast, this paper considers the influence of *changes* in bargaining strength and how reactions to these changes are influenced by repeated interaction.[7] Zwick and Chen (1999) examine a model similar to Binmore et al., in order to examine whether individuals take advantage of their bargaining strength, or if instead fairness concerns influence behavior. Their experiment also uses an alternating offer game where subjects pay costs, which could differ across players, for delays in reaching an agreement. Their experimental treatments vary the magnitude of the cost difference between the two players. Several other studies study similar games (Weg, Rapoport, and Felsenthal 1990; Anderhub, Guth, and Marchand 2004). None focus explicitly on the interaction between the probability of repeated interaction and asymmetric changes in bargaining strength, and how a larger shadow of the future can lead to commitment problems that have captured the attention of IR scholars.

The experimental literature on bargaining treats bargaining strength as static and often focuses on whether other-regarding preferences influence behavior.[8] The repeated games literature focuses on whether repetition can create incentives for cooperative behavior in places where single shot or finite play would not support cooperation. In contrast, I study the interaction between dynamic bargaining strength and the probability of repeated play. In the model I study, a

[6] ''This contrast between our one-shot and repeated play results is not encouraging news for those who might wish to interpret as gospel the oft-spoken suggestion that repeated play with discount factors close to one leads to more cooperative behavior. True enough, it does—but not by much'' (Palfrey and Rosenthal 1994:548).

[7] In the Binmore experiment, bargaining strength changes because the proposer position alternates and there is a shrinking pie, but this change is not asymmetric and does not create commitment problems for the current proposer.

[8] Furthermore, the bargaining literature typically investigates models where offer rejections are simply followed by another round of negotiation, possibly with a smaller total ''pie.'' The model considered in this paper has offer rejections followed by a costly lottery, where the winner gains control of the entire resource. This mimics resource division problems where suspension of negotiations and conquest by force can occur.

low probability of repeated play leads to a socially efficient outcome, whereas a high probability of repeated play leads to a socially inefficient outcome. The logic that gives rise to this result has nothing to do with mechanisms like concerns for fairness, spite, or reputation formation, the three central mechanisms explored in a recent review of ''Reasons for Conflict'' in bargaining games (Falk, Fehr, and Fischbacher 2003). That is, inefficient outcomes arise due to the incentives of the game, not psychological mechanisms or incomplete information about a player's type (as with reputation building).

Contributions to the Experimental and IR Literatures

While the shadow of the future figures centrally in the experimental and theoretical IR literatures, its empirical study is much less common. This appears to be due to difficulty in measuring the shadow of the future, understood either as how much leaders discount future utility or as the likelihood they expect to interact with an opponent in the future. For example, consider the recent comment by Ripsman and Levy (2007:42) in their discussion of the pre-World War II British and French appeasement of Germany: ''(t)radeoffs between current risks and future risks are sensitive to...time horizons, which vary across individuals *and which are extremely difficult to measure*...the time horizons of political leaders, particularly democratic political leaders, generally tend to be short, which reduces the incentives for current military actions in response to future threats.'' Levy (2008:21) makes a similar observation in his discussion of democracies and preventive war: ''Two key variables shaping leaders' threat perceptions and decisions are their time horizons and propensities toward risk-taking, though each is difficult to study empirically outside of a controlled laboratory setting.''[9] Finally, Fearon also notes the difficulty of studying the effect of discount rates:

> Nonetheless, we would like...to test for the specific and perhaps counterintuitive dynamic predicted by the war-of-attrition model that costly standoffs are more likely to occur in cases where state leaderships discount future payoffs relatively little. To do so, we need to be able to interpret and measure leaders' discount rates empirically, a difficult task since the number of factors that might influence a leadership's value for present versus future benefits is large (Fearon 1998:294).

Others have noted severe measurement challenges in the study of preventive war more generally (Lemke 2003).[10]

The laboratory experiment I report below makes a valuable contribution to the literature by carefully controlling and manipulating a variable that previous

[9] Levy goes on to make the interesting observation of Bismarck's change in mindset about preventive war over his tenure: ''Bismarck recognized both sides of the dilemma confronting policy makers. He stated that 'No government, if it regards war as inevitable even if it does not want it, would be so foolish as to leave to the enemy the choice of time and occasion and to wait for the moment which most convenient for the enemy.' Over time, however, Bismarck grew increasingly cautious. He said that 'preventive war is like suicide for fear of death,' and that 'We have to wait, rifle at rest, and see what smoke clouds and eruptions that volcano of Europe will bring forth' ''.

[10] Bearce, Floros, and McKibben (2009) attempt to measure the shadow of the future held by states and how this has an impact on their decision to bargain. They also note ''while the shadow of the future (our independent variable) has long been an important concept in international cooperation theory, scholars have not done much large-N empirical work using this concept'' (2009:726). They suggest that there will be a longer shadow of the future between two states when they are common members of more international governmental organizations, when they have higher rates of GDP growth, when governments do not face imminent reelections, and when elections are further away (2009:727–728). While their contribution is a welcome and important innovation, it is hard to argue that they cleanly measure the shadow of the future and not other phenomena. Fearon (1998:293) suggests that the shadow of the future should be long between parties with shared borders, between parties of the same country (civil conflict) and in issue areas with low exposure to exogenous shocks and broad regimes (as opposed to narrow agreements) (1998:295). Leventoglu and Tarar (2008:547) suggest the shadow of the future will be long when the amount of time between counteroffers is low.

studies have identified as causal yet have failed to measure with any precision. Furthermore, I study the effect of this manipulation in a strategic environment where the predicted relationships run the opposite direction from most existing work. If the behavioral test of the model is supported in this less-studied case, then we can be confident about the importance of the causal variable of interest, the shadow of the future. Laboratory experiments like the one in this paper help solve these challenges for empirically testing our theories. They also let us consider the extent to which an existing theoretical model is behaviorally realistic.

This study uses a game-theoretic model to generate hypotheses and design an experiment. Most work in experimental IR does not draw on game-theoretic reasoning, and if it does, the workhorse game is the widely studied prisoner's dilemma (for example, Majeski and Fricks 1995). Instead, IR scholars have used experiments to test non-strategic hypotheses (for example, Geva, Mayhay, and Skorick 2000; Wilkenfeld, Young, Asal, and Quinn 2003; Mintz 2004; Mintz, Redd, and Vedlitz 2006; Gates 2008). More recent work by IR scholars draws on game-theoretic reasoning and looks beyond experiments based on simple prisoner's dilemma games (for example, Butler, Bellman, and Kichiyev 2007; Tomz 2007; McDermott, Tingley, Cowden, Frazzetto, and Johnson 2009; Tingley and Wang 2010; Tingley and Walter 2011, forthcoming). The relative rarity of game-theoretically informed experimentation in the IR literature is perhaps surprising given the strong role strategic interaction plays in IR theories and the straightforward way many game-theoretic models translate into experimental designs.

Model

In this section, I summarize a simple game-theoretic model of repeated resource division proposed by Fearon (1995). Other models might also be used to study the negative relationship between the shadow of the future and social efficiency. I study this one and leave open the possibility for studying other models with experiments in the future. In this section, I characterize equilibrium behavior as a function of parameters in the game and then calculate a simple comparative static not previously considered that informs the experimental design of the next section. I define relative power (bargaining strength) as the commonly known probability of winning a lottery that assigns ownership of a contested resource to a single actor following the rejection of an offer.[11] Repeated interaction is modeled as a fixed and commonly known probability of continued interaction. Appendix 1 contains the relevant mathematical derivations.

Preliminaries

There are two actors, A and B, who interact repeatedly over multiple periods. Actor A proposes a division of the resource of size $R = 1$, $(x_{1A}, x_{1B} = 1 - x_{1A})$.[12] If B accepts the division, then both proceed to the second period. This procedure continues as long as the offer is never rejected. If the offer is rejected, then A and B incur costs $c_A$andc_B, respectively, to play a ''war lottery.'' With probability p_1, actor A wins the entire resource, and with probability $1 - p_1$, actor B wins the entire resource. The owner of the resource receives a payoff of R in every subsequent period and loser receives 0. The game is repeated and an additional period is reached with probability $\delta \in [0,1]$ known by both actors. With probability $1 - \delta$, the current period is the final period and there is no additional resource to be

[11] Such a lottery encapsulates what Fearon considers received wisdom that ''war is a gamble whose outcome may be determined by random or otherwise unforeseeable events'' (Fearon 1995:387). The model simplifies this series of events by representing the outcome as the result of a simple random variable.

[12] Numbers subscript periods and letters subscript actors.

divided or gained by the owner of the resource. If the war lottery is reached, there is no further interaction between A and B—the winner simply receives payoffs of size R in each additional period and the loser receives nothing.[13]

There is a key difference between period 1 and all future periods. The probability that actor A wins the war lottery increases by some amount from period 1, p_1, to period 2, p_2. If actors reach period 2, B decides whether to accept the offer or choose the war lottery as before. The game is infinitely repeated, but the shift in power only occurs between periods 1 and 2. All aspects of the game are common knowledge.[14]

Analysis

Fearon (1995) states that there exists an equilibrium under a set of parameter values where actor B will choose to launch a preventive war (rejects the offer) in period 1. The intuition is straightforward. Actor B knows that if period 2 is reached, he will receive a less favorable offer because of the shift in power. That is, in periods 2 and up, B expects to receive less than what he receives in the first period because A has a higher chance of winning the lottery. B weighs what he expects to get from offers in periods 2 and up and compares that to what he expects to get from rejecting the offer when he has a better bargaining position (period 1). Instead of accepting an offer in the first period, B will launch a preventive war (reject the offer in the first period). If A could somehow guarantee B more utility in periods 2 and up, this might be averted. Lacking such a mechanism, actor A faces a commitment problem. The change in power makes future exploitation more desirable, and A is unable to credibly commit to not exploiting his future power. This result of rejecting offers in the first period stands in contrast to the single play version of this resource division game, where, because of the costliness of the lottery, there always exists a resource division mutually preferred by both actors (assuming risk neutrality or aversion). *Hence, repeated interaction can lead to socially inefficient outcomes in equilibrium that are not present in a single period version of the game.* Whether B has an incentive to launch a preventive war depends on the change in bargaining strength (p_2,p_1), the cost to B of rejecting an offer (c_B), and the probability of continued interaction (the discount factor, δ). In particular, Fearon (1995) states the preventive war constraint as:

$$\delta p_2 - p_1 > c_b(1-\delta)^2 \tag{1}$$

When this inequality holds, any offer by A in the first period is rejected. Because all offers in equilibrium are rejected, A is indifferent over offers.

Preventive war does not happen when the inequality in Equation (1) does not hold. If Equation (1) does not hold, then in equilibrium we should not see first-period demands rejected. When this is the case, the equilibrium demand by A can be found:

[13] In a single-shot version of the game, the war lottery is never chosen because there exists a wedge of outcomes mutually preferred to war ($p - c_A, p + c_B$).

[14] The stage game here differs from other games used by experimentalists to study bargaining. First, the ''size'' of the pie does not shrink in each period. This resembles a situation where a resource is replenished, such as a piece of territory that is used for agriculture. Second, if an offer is rejected, then both actors pay a fixed cost. In IR, this cost is interpreted as a cost of war. Previous experiments instead use a shrinking pie to punish actors for rejected offers. Third, bargaining strength is implemented as the probability of gaining the entire resource following a rejected offer. In a single-shot setting, this would be equivalent to having an outside option (in expectation). However, the game is repeated and the actor gaining the resource continues to derive utility from it into the future. Fourth, rejected offers are followed by one actor gaining control of the resource. In an ultimatum game, a rejected offer leads to neither actor gaining anything which seems an unreasonable formulation if bargaining is occurring over territory or some other quantity of interest to IR scholars.

$$x_{1A}^{*} = \frac{p_1 - \delta p_2 + c_B(1-\delta)^2}{1-\delta} \tag{2}$$

Hence, Equation (2) gives an expression for the largest demand (1-offer) in the first period which will be accepted when the preventive war constraint does not hold. In the appendix, I re-derive the constraints on parameter values for preventive war, calculate comparative statics on the shadow of the future, and derive equilibrium offers under conditions where preventive war will not occur. I show that satisfaction of the preventive war constraint becomes more likely when the shadow of the future is longer, that is, as δ increases. Visual inspection of Equation (1) easily establishes this as well. This paper uses an experimental design to investigate this comparative static prediction. Furthermore, I show that as δ increases and Equation (1) does not hold (no preventive war), the demands by A in the first period are driven down. Intuitively, as the probability of reaching a period where actor A has a bargaining advantage becomes more likely, actor A will decrease his first-period demand to ensure that his offers in the first period are accepted. As this is an equilibrium relationship, actor B knows that the future is more valuable to actor A and expects to receive better offers in the first period.[15]

Experimental Design

In this section, I explain an experimental design that tests comparative static predictions about how changes in δ influence whether first-period rejections will occur and the size of first-period demands. A key advantage of experimental work based on a model is that the model helps guide the design of the experiment, but also helps generate hypotheses.

"Stage" Game

In the experiment, subjects are anonymously matched into pairs and play the game described above. One subject is randomly assigned to have position A and the other position B. Subject A then chooses an amount to offer subject B. Subject B sees the offer and chooses whether or not to reject the offer. If the offer is rejected, then a computer determines who gains the resources according to a commonly known probability and both pay commonly known costs $c_A = c_B$. Following the acceptance or rejection of an offer, the computer determines whether the pair will continue or will be terminated. This is done using a commonly known random stopping rule δ that I describe next.

Shadow of the Future

The key part of the experiment is the repeated interaction between subjects. Following others, I interpret the shadow of the future as the probability of future interaction.[16] To implement the shadow of the future in the laboratory, I use a probabilistic stopping rule (Dal Bo 2005). For example, if the discount factor is equal to 0.5, then in every period there is a 0.5 chance that there will be no future interaction. This probabilistic rule is formally equivalent to imposing a

[15] In the model above, the change in bargaining strength occurs only between periods 1 and 2. A reasonable question to ask then is what role the infinitely repeated setup has compared to a simpler two-period model. The answer is relatively simple. While there still exist values of $\delta, c_A, c_B, p_1, p_2$ such that a demand is rejected in the first period, the constraints are much tighter than in the infinitely repeated case. The parameters I use in the experiment are such that preventive war would not happen if the game were only repeated twice.

[16] Of course, other interpretations are possible. An alternative interpretation of discounting is through present valuations of future utility. This experiment does not directly manipulate this form of discounting and is limited to the probability of future interaction interpretation.

common discount factor on both subjects. If the war outcome is chosen (and costly lottery played), then the ''winner'' receives a sequence of payoffs equal to 1 until the stopping rule terminates the round. That is, if the war outcome is chosen, then the computer interface displays on the screen a number identifying the current period (1,2...) and the payoffs received until termination of the pair. If an offer is never rejected, then subjects continue to divide the resource until the interaction is terminated. When the stopping rule terminates play, subjects receive a message that says play has ended, lists their total payoffs (all periods added together), and asks them to wait while the computer pairs them with another opponent. In order to study the influence of the shadow of the future, one set of experiments uses a low probability of continued interaction and the other set uses a high probability.

Bargaining Strength

If the division was rejected in the first period, then the computer determines the owner, with A gaining it with probability p_1, B with probability $1 - p_1$, and both pay their respective costs c_A, c_B. They would then proceed to the second period with probability δ and have their interaction terminated with probability $1 - \delta$. If the period 1 offer was accepted, then period 2 is identical to period 1 except that the probability actor A gets the resource following a rejected offer changes such that $p_2 > p_1$. That is, between the first and second periods, there is a commonly known shift in power from p_1 to p_2. In each period, the computer displays on the screen the probability that subject A wins the war lottery.

Subject Rematching and Experiment Repetition

Once all subjects had their dyadic interactions terminated by the random stopping rule, subjects were rematched with someone that they had not yet faced. All matching was anonymous. After all subjects had been matched with every other subject, a *repetition* of the experiment was complete. Each experimental session had two repetitions with random matching, and subjects were paid based upon a randomly selected repetition of the experiment. Subjects were told that the experiment would repeat, but not the number of times that it would repeat, in order to minimize unmodeled meta-game effects. Repeating the experiment helps isolate learning effects that I document below. The literature on learning in experimental games is immense (for example, Cheung and Friedman 1997). I leave the modeling of learning in the experiment to future work and instead follow others by presenting results of the experiment broken out by subject experience with the game (Cheung and Friedman 1988; Ostrom, Walker, and Garnder 1992; Slonim and Roth 1998) but discuss hypothesis tests in reference to the final repetition and provide follow-up evidence from a separate experiment about learning.[17] Subjects were paid based on one randomly selected repetition of the experiment and received $1 for every point earned.

Parameter Selection

I use numerical values for parameters in the game that the model predicts should generate several key differences in observed behavior. The model suggests how changes in a key parameter, here δ, could influence behavior.[18] Sans

[17] Future work might use alternative designs to tease out differences between strategic adjustment and learning (Muller, Sefton, Steinberg, and Vesterlund 2008).

[18] This does not amount to using an experiment to merely test assumptions. Instead, I am testing implications of a particular assumption, isolating this part of the theory and manipulating it in an experiment.

TABLE 1. Model Parameters Used in Experiment

Model parameter	*Value in experiment*
Cost to player A for rejected offer, c_A	0.2
Cost to player B for rejected offer, c_B	0.2
Probability A wins resource if reject offer in period 1, p_1	0.3
Probability A wins resource if reject offer in period 2, p_2	0.7
Probability interaction continues (discount factor), δ	0.3 or 0.7

some explicit theoretical model, it would not be clear how to design an incentivized experiment that isolates the commitment problems of interest.

The experiment uses the following parameters summarized in Table 1. The costs for both subjects were the same, $c_A = c_B = 0.2$. The purpose of the experiment is not to isolate the role of differences in cost nor the role of uncertainty over costs. Hence, the costs of the two players are equal and commonly known.[19] A value of 0.2 represents 20% of the ''pie'' in a single period. This is a relatively low amount but still non-trivial.[20] I set the probability subject A wins the resource following an offer's rejection in the first period at $p_1 = 0.3$, which then shifts to $p_2 = 0.7$ in the second period and up. This represents a relatively large change on the probability scale and makes the commitment problem apparent but not necessarily self-evident. It is also symmetric around 0.5.[21] Given these parameters, I then selected two different values of δ that the model predicted should lead to differences in behavior. In the first treatment, I picked a high value, $\delta = 0.7$, which ensures that $\delta p_2 - p_1$ was sufficiently greater than $c_B(1 - \delta)^2$. In the second treatment, I chose a lower value that puts $c_B(1 - \delta)^2$ significantly lower than $\delta p_2 - p_1$, $\delta = 0.3$.[22]

Given these parameters, the model makes several specific predictions about what should occur in equilibrium. Following the experimental literature, I use these differences in predictions to state probabilistic hypotheses that I then test.

Hypothesis 1: *We should see more instances of rejected offers in the first period when there is a high shadow of the future* ($\delta = 0.7$) compared to when there is a low shadow of the future ($\delta = 0.3$).

Hypothesis 2: *Hypothesis 2 is a conditional version of Hypothesis 1. For a given offer or set of offers in period 1, the probability of rejection should be higher when* $\delta = 0.7$ than when $\delta = 0.3$. That is, taking Hypothesis 1 conditioning on actual offers.

Hypotheses 1 and 2 are the key hypotheses I examine. They concern the relationship between the likelihood of repeated interaction and the probability that socially efficient outcomes will obtain. Hypotheses 3a–3c test additional testable implications of the theory. The model pins down what demands should be when the preventive war constraint does not hold (Equation 2). When the constraint does hold, actor A is indifferent over demands, though Fearon suggests that A will make the lowest demand possible in this case. I thus pose several additional hypotheses.

[19] Future work using this design would easily be able to accommodate asymmetries and inject one- or two-sided uncertainty over the lottery cost.

[20] As a percentage of total expected earnings of both players, this poses a larger cost for those in the $\delta = 0.3$ case.

[21] Discussions in the design stage suggested little a priori reason to favor some other decision about a value to center the shift around.

[22] Furthermore, the set of parameters I chose were such that if the game were repeated twice, instead of being infinitely repeated, then equilibrium predictions were such that offers in the first period should not be rejected.

Hypothesis 3a: *In the low shadow of the future treatment* ($\delta = 0.3$), *demands from A should be* $x_{1A} = \frac{p_1 - \delta p_2 + c_b(1-\delta)^2}{1-\delta}$. *Given the parameters I use in the game, this equals = 0.27.*

Hypothesis 3b: *Following Fearon's heuristic example, those in the actor A position will offer a high amount in the first period when the preventive war constraint holds. As a result, offers in the first period will be higher when* $\delta = 0.7$ *compared to when* $\delta = 0.3$.

Hypothesis 3c: *If actor A is indifferent over his offers in the* $\delta = 0.7$ *treatment, then we should see a higher degree of variation in offers in period 1 compared to those in the* $\delta = 0.3$ *treatment.*

Results

In the spring of 2009, I recruited subjects from a set of students who had previously signed up to participate in experiments with a university's experimental social science laboratory.[23] Subjects were told that they would earn a $10 show-up fee plus any additional money that they earned during the experiment. Prior to participating in the experiment, subjects were not told anything additional about the experiment or the treatment condition to which they would be assigned. Throughout an experimental session, the values of δ, c_A, c_B, p_1, and p_2 were fixed and commonly known to all subjects. The treatment condition to be used for a given session was determined beforehand but again was unknown to subjects registering for the experiment. Because parameters were fixed within a given session, all analyses are between subjects.[24] Below, I present results from analyses of both repetitions of the experiment but focus on the second repetition—where subjects have the most experience with the game—in the text. In general, there was a greater difference between the treatments in the second repetition, which is likely due to learning. I discuss a follow-up experiment on learning below. These experimental practices conform to standard experimental game theory procedures. Covariate balance across treatments is reported in Appendix 2.

The experiment began with an instructional period where all features of the experiment were explained using oral instructions, written summaries of instructions distributed to subjects, and a slide presentation. Following the instructional session, subjects completed a practice session in order to become familiar with the computer interface.[25] After two practice rounds, subjects took a quiz on the key elements of the experiment. They could not proceed until they answered all questions correctly.

Following the experiment, subjects took a survey. A total of 36 subjects participated in each treatment, with 12 subjects in three sessions each. Average earnings in the $\delta = 0.3$ treatment were $17, and $24 when $\delta = 0.7$ for sessions lasting approximately an hour.

[23] It remains an open question as to whether results from student samples generalize to political leaders. One study by Mintz et al. (2006) suggests differences across samples in terms of the type of information acquired during a task with explicit substantive references to conflict situations. The current experiment does not provide subjects with any such substantive information so as to limit the extent to which subject differences could influence results depending on the substantive description of the experiment's task. Future work will use non-student samples but at present, there exist no clear reasons why non-student subjects should behave in a manner *opposite* that observed in the current study.

[24] It of course would be straightforward to conduct a within subject design.

[25] Computer interface programmed in Multistage, http://casselweb.ssel.caltech.edu.

Hypothesis 1

Hypothesis 1 states that we should see more instances of rejected offers in the first period when there is a higher shadow of the future ($\delta = 0.7$ versus $\delta = 0.3$). To test this hypothesis, I compare proportions of rejected offers in the first period across the two groups and use one-tailed tests given the directional hypotheses. Throughout, I cluster standard errors at the individual level.[26] Looking at the second repetition of the experiment, after which subjects had some experience with the strategic context, there was a significant difference in rejection rates across the treatments in the direction predicted by the model (51% when $\delta = 0.3$ and 63% when $\delta = 0.7$) ($N = 396$, $t = 1.37$, $p = .08$). Pooling offers of all sizes together I find reasonable support for Hypothesis 1 that a higher proportion of offers will be rejected, leading to both players paying a cost, when the shadow of the future was higher.

Hypothesis 2

Hypothesis 1 ignores the magnitude of the offer, which may in fact influence whether or not an offer is rejected. Hypothesis 2 takes Hypothesis 1 and conditions on the actual offers. I analyze Hypothesis 2 in several ways. First, I divide subjects by whether or not they received an offer below 0.5. When first-period offers were below 0.5, there was no significant difference in rejection rates across the treatments. Offers below the equal division threshold were rare and almost always rejected in both treatments. Again looking at the second repetition, the differences in rejection rates were sizable and highly significant (45% when $\delta = 0.3$ and 61% when $\delta = 0.7$) ($N = 362$, $t = 1.81$, $p < .05$). In the $\delta = 0.3$ treatment, the equilibrium offer, which should be accepted, is 0.73. Hence, the starkest differences between the two treatments should occur for offers over this amount. Because offers tended to take on rounded values like 0.7 and 0.8, here I restrict observations to first-period offers at 0.7 and above. Here, we see large and significant differences in rejection rates especially in the second repetition ($N = 294$, $t = 2.58$, $p < .01$). The complete set of results for Hypothesis 1 and 2 are given in Table 2.

Next, I estimated several probit models where I let the probability of rejection be a function of the offer and the treatment condition (1 if $\delta = 0.7$ and 0 if $\delta = 0.3$). Because the effect of offers on the probability of rejection could differ across treatments, I also allow for an interaction between δ and the offer. To illustrate the effect of δ, Figure 1 plots the estimated probability of rejection as a function of offers for the two treatment conditions. The top row uses all offers and responses. Because the distribution of offers tends to cluster above 0.5, the bottom row uses only observations with offers >0.5 (using only observations with offers at 0.7 and above produces similar results). The left panel uses both repetitions, the middle panel only the first repetition, and the right panel only the second repetition. The differential effect of offers across treatments is most apparent in the second repetition, where the probability of rejection declines much more quickly when $\delta = 0.3$ compared to when $\delta = 0.7$. This again suggests a strong learning dynamic. As subjects became more familiar with the strategic environment, they move closer to the equilibrium predictions.[27]

[26] Clustering at the individual level produces more conservative tests.

[27] Future research could derive a fully structural model of the game by extending results in Ramsay and Signorino (2010) to a dynamic repeated game setting. Generalized additive models (GAMs), which relax the parametric assumptions in the probit model, also give similar results.

TABLE 2. Difference in Proportion of Rejected-First Period Offers

Data	$\delta = 0.7(\%)$	$\delta = 0.3(\%)$	*p-value*
All offers			
Repetition 1 and 2	58	53	0.24
Repetition 1	54	55	0.38
Repetition 2	63	51	0.08
Offers ≥0.5			
Repetition 1 and 2	55	42	0.05
Repetition 1	48	38	0.15
Repetition 2	61	45	<0.05
Offers ≥0.7			
Repetition 1 and 2	53	31	<0.01
Repetition 1	43	24	<0.05
Repetition 2	61	35	<0.01

(*Notes*. Difference in proportion of rejected first-period offers. Standard errors clustered at individual level and *p*-values from one-sided tests.)

Hypothesis 3a

When there is a low shadow of the future and the preventive war constraint does not hold, the model predicts offers of 0.73. To test this prediction, I calculate average offers for the $\delta = 0.3$ case and then ask whether the average offer differs significantly from 0.73. I do not find support for the point prediction in Hypothesis 3a. The average offers when $\delta = 0.3$ were regularly lower than 0.73. However, while the average offer in repetition 1 was 0.56, this increased to 0.63 in repetition 2. Thus, while offers were not close to the point prediction made by the model, the direction of learning was toward the equilibrium.

Hypothesis 3b

Hypothesis 3b suggests that the offer in period 1 will be higher when $\delta = 0.7$ compared to offers when $\delta = 0.3$. Figure 2 displays the distribution of offers across the two treatments. The vertical lines represent the means of the two samples. A range of statistical tests reveal differences between the two distributions.[28] Offers under a high shadow of the future were significantly greater in comparison with the low shadow of the future treatment. While the model predicts multiple equilibria with offers anywhere along [0,1] when $\delta = 0.7$, empirically we observe average offers that are higher compared to the $\delta = 0.3$ case. This is somewhat consistent with Fearon's heuristic example of the offer in the first period being the entire pie when Equation (1) holds (that is, in the $\delta = 0.7$ case). As before, the magnitude of the difference between the treatments is largest in the second repetition. This again illustrates a learning dynamic in the game.

Hypothesis 3c

Hypothesis 3c addresses the result from the model that when Equation (1) holds, all offers will be rejected and actor A is indifferent over what offer to make in the first period. One possible interpretation of this indifference is that offers will be more spread out compared to when Equation (1) does not hold. That is, offers will have higher variance when $\delta = 0.7$. There is little evidence of this, as can be seen from Figure 2. While the average offer is significantly

[28] I calculated difference in means tests, difference in means with standard errors allowing for correlated errors within subjects, and Wilcoxon rank-sum tests. All suggest highly significant differences.

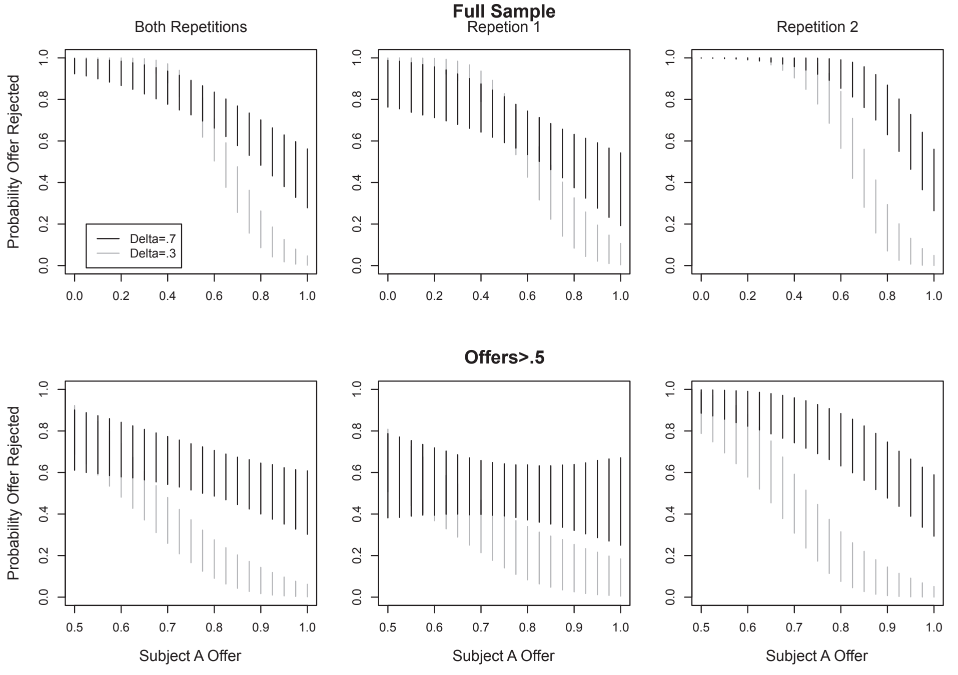

Fig 1. Probability of Offer Rejection

(*Notes:* Top row reports models using all offers, and bottom row estimated using only observations where the offer was >0.5. Left panel estimated from probit model with offer size, treatment condition, interaction between offer and treatment, and dichotomous variable for the repetition of the experiment. Middle panel estimated using only the first repetition, and right panel using only the second repetition. Estimated probability of rejection for $\delta = 0.7$ with 95% confidence intervals displayed in black. Estimated probability of rejection when $\delta = 0.3$ in grey. Standard errors clustered at individual level and confidence intervals calculated using a parametric bootstrap running for 1000 iterations. Each figure displays a positive interaction between the probability of rejection and being in the treatment with a high shadow of the future. The interaction is strongest in repetition 2.)

higher when $\delta = 0.7$, there does not appear to be greater variation in offers. The differences in variances were not significant according to Levene (1960) robust test of difference in variances. While there is variation in the offers when $\delta = 0.7$, the perceived incentive to make a high offer makes this variability relatively lower.

Post-experiment Strategies

In a post-experiment survey, I asked all subjects two questions to solicit the strategies they used in the experiment: (i) ‘‘Think back to the experiment you just participated in. Imagine you were in the B position and in the first period. What was the MINIMUM offer you had to receive in order to ACCEPT the offer? Please fill in a value between 0 and 1.’’ (Several subjects said they would accept no offer and these cases are coded as a 1 for presentational purposes.) (ii) ‘‘Imagine you were in the A position and in the first period. What was the MINIMUM offer you felt you had to make in order for the OTHER person to accept the offer? Please fill in a value between 0 and 1.’’ These questions solicit the basic strategies relevant to Hypotheses 1 and 2 that subjects formulated after participating in the experiment. Figure 3 plots each of these quantities for the two treatment conditions. The significant differences observed in actual play,

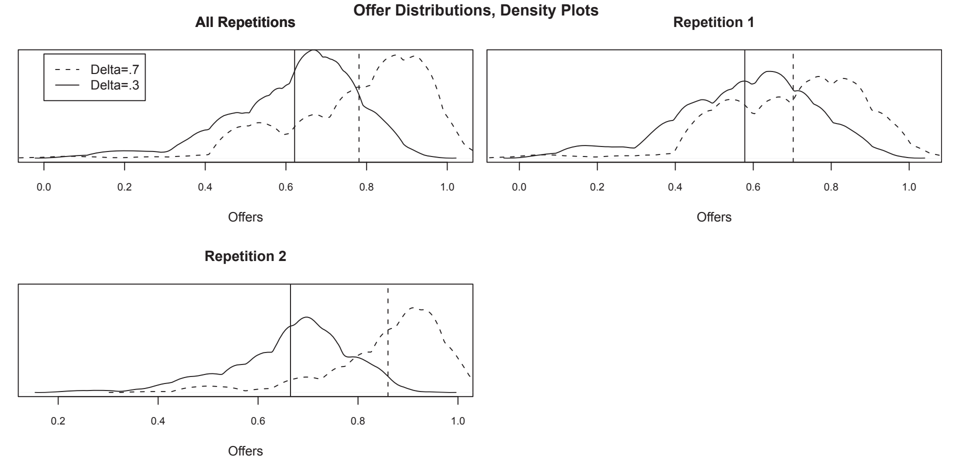

Fig 2. Distribution of Offers by Treatment Condition
(*Notes:* Plots are kernel density estimates with vertical lines indicating means. Top left panel pools all repetitions together. Top right panel includes only the first repetition and bottom left panel includes only the second repetition. In each case, there are statistically significant differences in offers across the treatment conditions, with higher offers in the $\delta = 0.7$ case.)

especially in the second repetition, are also reflected in the post-experiment reported strategies. Offers were higher (Wilcoxon ranksum: $N = 72$, $z = -3.78$, $p < .01$) and rejection threshold higher in the first period ($N = 72$, $z = -4.5$, $p < .01$) in the $\delta = 0.7$ treatment. These results provide additional support for the key hypotheses of the paper.

An Additional Look at Learning

In Table 2 and Figure 1, there was evidence of learning across the two repetitions of the experiment. As subjects had more experience with the experiment, the differences across the treatments became larger. To investigate this dynamic further, I offer a final piece of evidence that involves a separate sample of students enrolled in an undergraduate game theory class taught in a political science department. This lecture class held separate meetings with smaller numbers of students to review material each week. One of the final topics covered in class was a section on repeated games. The review section meetings held prior to the introductory lecture participated in the above experiment before the students heard the lecture.[29] The other half of review sections participated in the experiment after the lecture. For both groups, I use the high shadow of the future treatment ($\delta = 0.7$).[30] I ask two questions: after subjects had gone through an introductory lecture on repeated games (i) were offer rejections in the first period more likely and (ii) were first-period offers higher?

[29] The introductory lecture on repeated games covered basic concepts about repeated play, discounting, and the use of conditional punishment strategies in a PD game.

[30] All features of the experiment were the same except for the following changes that were logistically necessary. (i) Subjects were paid using extra credit points for the class instead of money (regulations prevented the use of money). (ii) Review sections had different numbers of individuals and so we simply had subjects paired up with the same number of persons in the main experiment (11) but because sections had less than this number of people, this meant people were paired with each other more than once. However, as all pairing was anonymous, subjects did not have the ability to link past behavior to current decision making. In these sessions, the experiment was repeated only once, after 11 sets of pairing the experiment was over, because of time constraints.

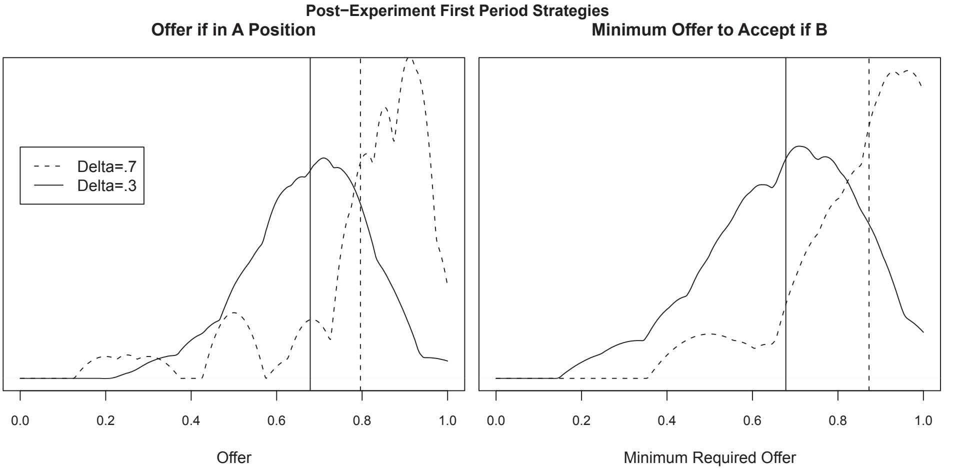

FIG 3. Post-experiment Solicited Strategies
(*Notes:* Distribution of offers and minimum required offers by treatment condition. Plots are kernel density estimates with vertical lines indicating means. Left panel plots offers if in the A position, and the right panel plots the minimum required offer in the B position. In each case, there are statistically significant differences in offers across the treatment conditions, with higher offers and required offers in the $\delta = 0.7$ case.)

To answer the first question, I estimated the probability of offer rejection in the first period as a function of offers and whether someone took the experiment before or after the lecture (using the same estimator as in Figure 1 and setting offers to the mean). Those who took the experiment after the lecture were 9% (95% CI: −5%, 25%) more likely to reject offers in the first period. To answer the second question, I compare the distribution of offers across the two groups. Figure 4 plots these distributions and their respective means. Difference in means tests with standard errors clustered at the individual level reveal this to be a significant difference ($N = 335, t = 1.85, p < .1$). In both cases, exposure to ideas about the influence of the shadow of the future on rational strategic

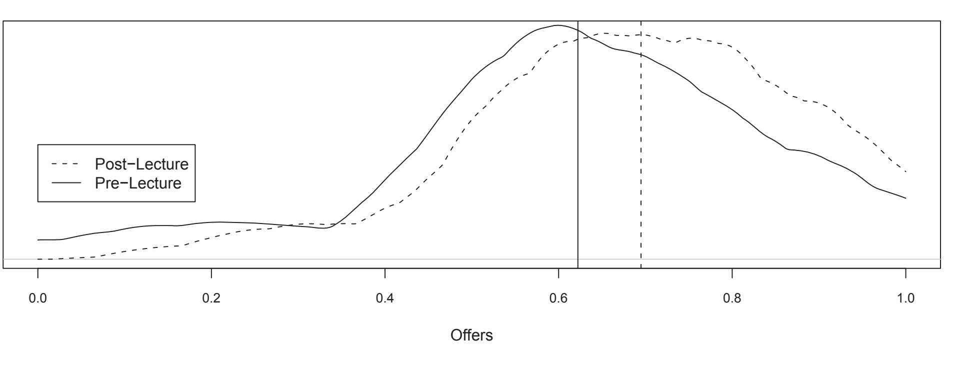

FIG 4. Pre- and Post-Lecture Offer Distributions
(*Notes:* Distribution of offers prior to lecture versus distribution of offers following lecture. Plots are kernel density estimates with vertical lines indicating means. Offers were significantly higher post-lecture.)

behavior led to changes in behavior similar to those observed across repetitions in the main experimental sessions. Further, subjects on average did not seem to engage in analogical reasoning (Jervis 1976) by simply trying to cooperate more, as the lesson on the repeated prisoner's dilemma might suggest. A better understanding of how repeated play influences strategic interaction—a characteristic likely in populations of decision makers outside the laboratory—leads to decisions closer to the equilibrium predictions.

Conclusion

In this paper, I study the interaction between commitment problems in bargaining and the likelihood of continued interaction. A longer ''shadow of the future'' does not necessarily lead to an increasing chance of socially efficient outcomes. In fact, it can have the opposite effect: a ''dark side of the future.'' Using incentivized laboratory experiments, I find substantial support for the model's key prediction of more rejected offers in the first period—more preventive war—when there is a longer shadow of the future. Using a game theoretic model, I pin down a set of parameters that allow me to make a simple comparison of behavior when the probability of continued interaction is low ($\delta = 0.3$) versus high ($\delta = 0.7$). The difference across the treatments is larger in the second repetition of the experiment, when subjects had more experience with the game. Furthermore, the difference across treatments holds even when I control for offers. Subjects in the B position clearly appear to anticipate a substantial change in offer if period 2 is reached and are more prone to reject offers when $\delta = 0.7$. A longer shadow of the future magnifies the commitment problem created by changes in bargaining strength. These results speak both to the literature on commitment problems due to shifts in power (for example, Powell 2006) and the effect of repeated play on the shadow of the future (for example, Oye 1985).

The results in this paper stand in contrast to the usual mantra on repeated interaction and a long shadow of the future. The IR and experimental literatures typically draw on a model akin to the repeated prisoner's dilemma and show how a low probability of continued interaction leads to socially inefficient behavior whereas increasing the probability of future interaction can lead to increases in the frequency of socially efficient outcomes. This picture of the shadow of the future looks very different from the one I examine. Nevertheless, this paper shares an important similarity with earlier work on repeated play. I identify substantial differences in behavior due to the effect of changes in the probability of continued interaction. I thus provide new evidence on the effect of repeated play on strategic behavior in a game of long-standing interest to IR scholars. I also help fill in the lacuna of work that several scholars have noted (for example, Fearon 1998; Ripsman and Levy 2007; Toft 2007; Levy 2008) on the effect of the shadow of the future in dynamic bargaining environments. To do this, I use laboratory experiments that offer one way to overcome the measurement challenges noted by these authors. Hence, the contributions of the paper are to test a canonical model in IR, analyze a comparative static prediction difficult to study with field data, provide evidence challenging a conventional wisdom about the effect of repeated play, and bring a new design to the experimental literature on bargaining motivated by IR questions.

There are several avenues for future research. This paper focuses on one comparative static to the model. Changing the magnitude of changes in bargaining strength, for a given value of δ, would let us test other predictions. This would allow a more complete investigation of the interaction between changes in bargaining strength and the probability of continued interaction. Another approach would be to vary the magnitude of the cost terms, or even have those parameters be private information as is commonly assumed in informational accounts of conflict. Another change would be to move beyond student samples. For example, military officers (for example, Mintz et al., 2006) might respond in different ways.[31] Similarly, the substantive description of the experiment could be manipulated to include labels on the choice variables and identities to the roles played by subjects (for example, ethnic group X or Y). This type of design would use strategic experiments as ''scaffolding'' for investing the impact of identity politics (Abdelal, Herrera, Johnston, and McCermoot 2006; Habyarimana, Humphreys, Posner, and Weinstein 2009). All of these are straightforward modifications to the experimental design and might minimize externally validity concerns.

Second, I hope to evaluate various ways to lessen the commitment problem created by future changes in bargaining strength. Chadefaux (2009) shows that allowing players to bargain over the distribution of power can eliminate the commitment problem. Eilstrup-Sangiovanni and Verdier (2005) argue that recourse to multilateral institutions can help solve commitment problems. Could ''cheap talk'' between opponents resolve the commitment problem, or are costly mechanisms required given the opposing interests of the players? Laboratory tests of these and other solutions would enable stress tests of various institutional designs. Hence, this type of work could play an important role in designing and evaluating policy interventions that confront commitment problems, such as in civil war termination (Walter 1999). Similarly, Fearon, Maccarten and Weinstein (2009) consider the relationship between policy interventions designed to reduce conflict and levels of social cohesion. They measure the effect of policy interventions on local members of the population through play in prisoner's dilemma-type games. These new approaches to studying the effect of policy interventions on violence might consider alternative strategic dilemmas such as the one this paper considers. If commitment problems underlying conflict involve shifts in power (Powell 2006), then behavioral measures of policy intervention/peace building efforts should capture their effects and examine their relationship to social cohesion. Using games like the one studied here will give field researchers a more robust experimental toolbox for studying conflict.

Appendix 1: Equilibrium Analysis

Preventive War

Fearon shows the unique sub-game perfect equilibrium offer in the second period is $x_{2A} = p_2 + c_B(1 - \delta)$.[32] This offer makes actor B indifferent in period 2 between war and accepting. This is straightforward to see:

$$\begin{aligned} \frac{1 - p_2}{1 - \delta} - c_B &= \frac{1 - x_{2A}}{1 - \delta} \\ 1 - p_2 - c_B(1 - \delta) &= 1 - x_{2A} \\ x_{2A} &= p_2 + c_B(1 - \delta) \end{aligned}$$

[31] An elite sample might have more experience in situations with shifting power and hence behavior would converge to equilibrium behavior earlier, in which case such an extension would find more behavioral support for the model. Crucially, however, it is not clear why an elite sample would be less likely to be effected by changes in δ. One difference might be that a non-student sample would likely have a wider distribution of political ideology (college students in my sample tended to be more liberal). Previous work suggests that political ideology is an important predictor of support for preventive war, with more conservative individuals supporting preventive action more (Silverstone 2007:189). Hence, one prediction would be that non-student samples might engage in more preventive war choices in the experiment, but there is no reason that such differences should interact with δ in a way that would suggest bias in the comparative static results.

[32] I assume that when actor B is indifferent he accepts the demand.

Because p_2 and c_B remain fixed for every additional period, the accepted equilibrium offer will remain the same in periods two and up. The backward induction solution starting at the nth period will have the same strategies played as the second stage.

For this particular demand to be in equilibrium in the second period, it must also be the case that actor A prefers making it compared to some other higher demand. Per above, any other higher demand would lead to actor B choosing the lottery. Hence, actor A compares the present discounted value of demanding x_{2A} in period 2 and all future periods to the present discounted value of the lottery. If x_{2A} is accepted forever, then we have $\frac{p_2 + c_B(1-\delta)}{1-\delta}$.The present discounted value for war is $\frac{p_2}{1-\delta} - c_A$. The offer x_{2A} will be as least as good as war if $\frac{p_2 + c_B(1-\delta)}{1-\delta} \geq \frac{p_2}{1-\delta} - c_A$, which holds because $c_A, c_B > 0$. Hence, actor A will always prefer acceptance of his demand $x_{2A} = p_2 + c_B(1 - \delta)$.

Now consider the demand in the first period. Fearon's exposition proceeds in a way that makes rejection of a demand as unlikely as possible. His purpose was to show how preventive war was possible, and hence, he uses as a heuristic the case where an offer being rejected in the first period is the largest possible offer. The best that actor A can do for actor B in order to avoid war in the first period is to give him the entire resource, $x_{1A} = 0, x_{1B} = 1 - x_{1A} = 0$. In the second period, actor B will get the offer from actor A equal to $1 - x_{2A}$, which makes state B indifferent and accepting by definition. Thus, actor B will go to war in the first period when:

$$\frac{1-p_1}{1-\delta} - c_B > 1 + \frac{\delta(1-x_{2A})}{1-\delta} \tag{3}$$

Substituting in for the optimal x_{2A} yields $\frac{1-p_1}{1-\delta} - c_B > 1 + \frac{\delta(1-[p_2+c_B(1-\delta)])}{1-\delta}$ which can be reduced to what Fearon states as his preventive war constraint, Equation (1): $\delta p_2 - p_1 > c_B(1-\delta)^2$

This constraint is derived under the assumption that actor A demands $x_{1A} = 0$. If the demand were greater than this, $x_{1A} > 0$, it is straightforward to show that Equation (1) still holds. Hence, no matter the offer in the first period, if Equation (1) holds, then the offer will be rejected. In equilibrium, it is not the case that actor A will necessarily offer the whole resource to B in the first period. Instead, there exist a continuum of equilibria when $\delta p_2 - p_1 > c_B(1-\delta)^2$ where actor A demands $x_{1A} \in [0,1]$, $x_{2A} = p_2 + c_B(1-\delta)$, and actor B rejects all demands in the first period.

No Preventive War

Now consider the case where preventive war does not occur. The second-period offer remains the same as above, $x_{2A} = p_2 + c_B(1-\delta)$. Actor A's present discounted value for war in the first-period is $\frac{p_1}{1-\delta} - c_A$. If the offer is accepted, then he receives $x_{1A} + \delta V_2$ where V_2 is the continuation value of the game starting at period 2. From above, his demand will be $p_2 + c_B(1-\delta)$ in period 2 and up. I assume that when actor B is indifferent, he accepts the demand, and hence we can substitute for $V_2 = \frac{x_{2A}}{1-\delta} = \frac{p_2 + c_B(1-\delta)}{1-\delta}$. Actor A will prefer having the demand x_{1A} accepted to war when:

$$x_{1A} + \frac{\delta x_{2A}}{1-\delta} \geq \frac{p_1}{1-\delta} - c_A \tag{4}$$

$$x_{1A} \geq \frac{-\delta p_2 - \delta c_B(1-\delta) + p_1}{1-\delta} - c_A \tag{5}$$

Now find x^*_{1A} such that actor B is indifferent between accepting the first-period offer and selecting war in the first period:

$$1 - x^*_{1A} + \frac{\delta(1 - x_{2A})}{1 - \delta} = \frac{1 - p_1}{1 - \delta} - c_b \tag{6}$$

$$\frac{p_1 - \delta p_2 + c_b(1 - \delta)^2}{1 - \delta} = x^*_{1A} \tag{7}$$

Hence, Equation (2) (replicated above) gives an expression for the largest demand in the first period, which will be accepted. Specifically, any demand greater than this will generate preventive war. B will accept lower demands (down to $x'_{1A} = \frac{-\delta p_2 - \delta c_B(1-\delta) + p_1}{1-\delta} - c_A$) but A has no incentive to make them.

Now recall that $x^*_{1A} \in [0, 1]$. Under what conditions will Equation (2) be <0?

$$\frac{p_1 - \delta p_2 + c_B(1 - \delta)^2}{1 - \delta} < 0$$

$$\delta p_2 - p_1 > c_B(1 - \delta)^2$$

This is exactly the preventive war constraint in equation (1). In other words, when these conditions hold, actor A cannot make a generous enough offer in order to avoid preventive war. This is the heart of the commitment problem.

Comparative Statics

Fearon states that as the magnitude of $p_2 - p_1$ increases, the preventive war constraint is more likely to hold. This is clear from Equation (1), but Fearon does not discuss how the discount factor changes whether this condition is met. The influence of changing δ is the key focus of this paper. How does changing the probability of future interaction change behavior in the first period of interaction?

Visual inspection of $\delta p_2 - p_1 > c_b(1 - \delta)^2$ clearly shows the hypothesized relations. This can be verified simply. Collect terms from $\delta p_2 - p_1 > c_b(1 - \delta)^2$ on the left-hand side and differentiate with respect to δ.

$$\frac{\partial}{\partial \delta}[\delta p_2 - p_1 - c_b(1 - \delta)^2] = p_2 + 2c_b(1 - \delta) \tag{8}$$

Here, $p_2, c_b > 0$ and $(1 - \delta) > 0$. Thus, the function on the right-hand side is positive. As δ increases, there is a threshold where actor B values the future enough to warrant their wanting to wage a preventive war and choose the lottery in the first period. Furthermore, the range of $\delta' s$ for which a preventive war will be fought is increasing in p_2. For fixed values of δ, p_1, c_b, increases in p_2 make the preventive war constraint more likely to be satisfied. That is, with larger power shifts, δ can be lower and preventive war will still occur.

The experiment has a condition where the preventive war constraint is not met ($\delta = 0.3$). While Fearon does not discuss what demands should be in this case, it is straightforward to solve for them and show how they change with δ. The equilibrium demand by actor A in the first period is $x^*_{1A} = \frac{p_1 - \delta p_2 + c_b(1-\delta)^2}{1-\delta}$. Differentiation with respect to δ yields:

$$\frac{d}{d\delta}\left[\frac{p_1 - \delta p_2 + c_b(1-\delta)^2}{1-\delta}\right] =$$

$$\frac{d}{d\delta}\left[1 + \frac{\delta - \delta p_2 - \delta c_b + \delta^2 c_b - 1 + p_1}{1-\delta} + c_B\right] =$$

$$= \frac{-p_2 - c_b + 2\delta c_b - \delta^2 c_b + p_1}{(1-\delta)^2}$$

$$= \frac{p_1 - p_2 - c_b(\delta - 1)^2}{(1-\delta)^2}$$

Now recall that $p_1 - p_2 < 0$ ($p_1 < p_2$) and $-c_B(\delta - 1)^2 \leq 0$. Thus, the derivative is <0. Intuitively, as the period of time that actor A has a bargaining advantage becomes more valuable, actor A will decrease his first-period demand to ensure that his offer in the first period is accepted. As this is an equilibrium relationship, actor B knows that the future is more valuable to actor A and expects to receive better offers in the first period.

Appendix 2: Experiment Descriptive Statistics

Experiment participants were recruited from an existing pool of subjects that had signed up with the university's social science laboratory subject pool. Upon signing up, subjects were not told anything about the experiment or the treatment condition in which they would participate, thus limiting potential for non-random assignment to treatment conditions. In each experimental session, there were 12 subjects, and for each experimental treatment condition, there were three experimental sessions. Across several subject-specific characteristics measured after the experiment, there existed no significant differences between the two treatment conditions. Means and standard errors by treatment condition are reported in Table 3 and refer to variables defined in the text. The multivariate omnibus balance check described in Hansen and Bowers (2008) produces an insignificant $p = .36$.

TABLE 3. Average Subject and Experiment Characteristics by Treatment Condition

	Treatment $\delta = 0.3$	*Condition* $\delta = 0.7$
Gender	0.5 (0.09)	0.45 (0.08)
Ideology	2 (0.25)	2.3 (0.25)
Economics class	0.64 (0.08)	0.69 (0.08)
Age	21.2 (0.52)	20.8 (0.24)
Average number of periods per dyad	1.45	3.57
Average number of offers per dyad	1.19	1.71

(*Notes:* Means with standard errors in parentheses. *Ideology* is a standard seven point measure of liberal/conservative political attitudes taken from the World Values Survey (when it comes to politics, do you usually think of yourself as Very liberal....Very conservative). *Gender* is coded as 1 for male subjects and 0 for females. *EconClass* codes whether or not someone had taken an economics class.)

References

ABDELAL, R., Y. HERRERA, A. I. JOHNSTON, AND R. MCCERMOTT. (2006) Idendity as a Variable. *Perspectives on Politics* 4 (4): 695–711.

ANDERHUB, V., W. W. GUTH, AND N. MARCHAND. (2004) Early or Late Conflict Settlement in a Variety of Games—An Experimental Study. *Journal of Economic Psychology* 25 (2): 177–194.

Andreoni, J., and J. H. Miller. (1993) Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence. *The Economic Journal* 103 (418): 570–585.

Axelrod, R. (1984) *The Evolution of Cooperation.* New York: Basic Books.

Axelrod, R., and R. O. Keohane. (1985) Achieving Cooperation Under Anarchy: Strategies and Institutions. *World Politics* 38 (1): 226–254.

Barkin, S. (2004) Time Horizons and Multilateral Enforcement in International Cooperation. *Journal of Conflict Resolution* 48 (2): 363–382.

Bearce, D., K. Floros, and H. McKibben. (2009) The Shadow of the Future and International Bargaining: The Occurence of Bargaining in a Three-Phase Cooperation Framework. *Journal of Politics* 71 (2): 719–732.

Binmore, K., P. Morgan, A. Shaked, and J. Sutton. (1991) Do People Exploit Their Bargaining Power? An Experimental Study. *Games and Economic Behavior* 3 (3): 295–322.

Butler, C. K., M. J. Bellman, and O. A. Kichiyev. (2007) Assessing Power in Spatial Bargaining: When Is There Advantage to Being Status-Quo Advantaged? *International Studies Quarterly* 51 (3): 607–623.

Chadefaux, T. (2009) *"Bargaining Over Power: When Do Rapid Shifts in Power Lead to War?"* Working paper. Available at http://www-personal.umich.edu/~chadefau/BoP.pdf. (Accessed March 19, 2011.)

Cheung, Y.-W., and D. Friedman. (1988) Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism. *Quarterly Journal of Economics* 103 (1): 179–200.

Cheung, Y.-W., and D. Friedman. (1997) Individual Learning in Normal Form Games: Some Laboratory Results. *Games and Economic Behavior* 19 (1): 46–76.

Dal Bo, P. (2005) Cooperation Under the Shadow of the Future Experimental Evidence from Infinitely Repeated Games. *The American Economic Review* 95 (5): 1591–1604.

Duffy, J., and J. Ochs. (2009) Cooperative Behavior and the Frequency of Social Interaction. *Games and Economic Behavior* 66 (2): 785–812.

Eilstrup-Sangiovanni, M., and D. Verdier. (2005) European Integration as a Solution to War. *European Journal of International Relations* 11 (1): 99–135.

Falk, A., E. Fehr, and U. Fischbacher. (2003) Reasons for Conflict: Lessons from Bargaining Experiments. *Institutional and Theoretical Economics* 159 (1): 171–195.

Fearon, J. (1995) Rationalist Explanations for War. *International Organization* 49 (3): 379–414.

Fearon, J. (1998) Bargaining, Enforcement, and International Cooperation. *International Organization* 52 (2): 269–305.

Fearon, J., and D. Laitin. (2008) *"Civil War Termination".* Working paper. Available at http://www.stanford.edu/~jfearon/papers/termination.pdf. (Accessed March 19, 2011.)

Fearon, J., H. Maccarten, and J. Weinstein. (2009) Can Development Aid Contribute to Social Cohesion After Civil War? Evidence from a Field Experiment in Post-Conflict Liberia. *American Economics Reveiw: Papers and Proceedings* 99 (2): 287–291.

Fischer, F. (1961) *Germany's Aims in the First World War.* New York: W.W. Norton.

Garfinkel, M. R., and S. Skaperdas. (2000) Conflict Without Misperceptions or Incomplete Information: How the Future Matters. *Journal of Conflict Resolution* 44 (6): 793–807.

Gartzke, E. (1999) War Is in the Error Term. *International Organization* 53 (3): 567–587.

Gates, S. (2008) The Multiple Effects of Casualties on Public Support for War: An Experimental Approach. *The American Political Science Review* 102 (1): 95–106.

Geva, N., J. Mayhay, and M. Skorick. (2000) The Cognitive Calculus of Foreign Policy Decision Making: An Experimental Assessment. *Journal of Conflict Resolution* 44 (4): 447–471.

Habyarimana, J., M. Humphreys, D. Posner, and J. Weinstein. (2009) Why Does Ethnic Diversity Undermine Public Goods Provision? *Amerian Political Science Review* 101 (4): 709–725.

Hansen, B., and J. Bowers. (2008) Covariate Balance in Simple, Stratified and Clustered Comparative Studies. *Statistical Science* 23 (2): 219–236.

Hess, G., and A. Orphanides. (1995) War Politics: An Economic, Rational-Voter Framework. *American Economic Review* 85 (4): 828–846.

Jervis, R. (1976) *Perception and Misperception in International Politics.* Princeton, NJ: Princeton University Press.

Keohane, R. (1986) Reciprocity in International Relations. *International Organization* 40 (1): 1–27.

Kim, W., and J. D. Morrow. (1992) When Do Power Shifts Lead to War? *American Journal of Political Science* 36 (4): 896–922.

Lemke, D. (2002) *Regions of War and Peace.* Cambridge, UK: Cambridge University Press.

Lemke, D. (2003) Investigating the Preventive Motive for War. *International Interactions* 29 (4): 273–292.

Levene, H. (1960) Robust Tests for Equality of Variances. In *Contributions of Probability and Statistics: Essays in Honor of Harold Hotelling*, edited by Ingram Olkin. PaloAlto, CA: Stanford University Press.

Leventoglu, B., and A. Tarar. (2008) Does Private Information Lead to Delay or War in Crisis Bargaining? *International Studies Quarterly* 52: 533–553.

Levy, J. (1987) Declining Power and the Preventive Motivation for War. *World Politics* 40 (1): 82–107.

Levy, J. (2008) Preventive War and Democratic Politics. *International Studies Quarterly* 52 (1): 1–24.

Majeski, S., and S. Fricks. (1995) Conflict and Cooperation in International Relations. *The Journal of Conflict Resolution* 39 (4): 622–645.

McBride, M., and S. Skaperdas. (2009) ''Conflict, Settlement, and the Shadow of the Future''. CESifo Working Paper Series No. 2897. Available at http://ideas.repec.org/p/ces/ceswps/_2897.html#provider (Accessed March 19, 2011.)

McDermott, R., D. Tingley, J. Cowden, G. Frazzetto, and D. Johnson. (2009) Monoamine Oxidase A Gene (MAOA) Predicts Behavioral Aggression Following Provocation. *Proceedings of the National Academy of Sciences* 106 (7): 980–994.

Mintz, A. (2004). Foreign Policy Decision Making in Familiar and Unfamiliar Settings: An Experimental Study of High-Ranking Military Officers. *Journal of Conflict Resolution* 48 (91): 91–104.

Mintz, A., S. B. Redd, and A. Vedlitz. (2006) Can We Generalize from Student Experiments to the Real World in Political Science, Military Affairs, and International Relations? *Journal of Conflict Resolution* 50 (5): 757–776.

Morton, R., and W. Williams. (2010) *From Nature to the Lab: Experimental Political Science and the Study of Causality*. Cambridge, UK: Cambridge University Press.

Muller, L., M. Sefton, R. Steinberg, and L. Vesterlund. (2008) Strategic Behavior and Learning in Repeated Voluntary Contribution Experiments. *Journal of Economic Behavior and Organization* 67 (3–4): 782–793.

Olson, M. (2000). *Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships*. New York: Basic Books.

Ostrom, E., J. Walker, and R. Gardner. (1992) Covenants With and Without a Sword: Self-Governance Is Possible. *The American Political Science Review* 86 (2): 404–417.

Oye, K. (1985). Explaining Cooperation Under Anarchy: Hypotheses and Strategies. *World Politics* 38 (1): 1–24.

Palfrey, T., and H. Rosenthal. (1994) Repeated Play, Cooperation, and Coordination: An Experimental Study. *Review of Economic Studies* 61 (3): 545–565.

Powell, R. (1996) Uncertainty, Shifting Power, and Appeasement. *American Political Science Review* 90 (4): 749–764.

Powell, R. (1999) In *the Shadow of Power: States and Strategies in International Politics*. Princeton, NJ: Princeton University Press.

Powell, R. (2006) War as a Commitment Problem. *International Organization* 60: 169–203.

Ramsay, K., and C. Signorino, (2010) ''A Statistical Model of the Divide-the-Dollar Game''. Working paper. Available at http://www.princeton.edu/~kramsay/Site/research_files/mechdesign3.1.pdf. (Accessed March 19, 2011.)

Rapkin, D., and W. Thompson. (2003) Power Transition, Challenge, and the (Re)Emergence of China. *International Interactions* 29 (4): 315–342.

Ripsman, N., and J. Levy. (2007) The Preventive War that Never Happened: Britain, France, and the Rise of Germany in the 1930s. *Security Studies* 16 (1): 32–67.

Roth, A., and J. Murnighan. (1978) Equilibrium Behavior and Repeated Play of the Prisoners' Dilemma. *Journal of Mathematical Psychology* 17 (2): 189–198.

Schweller, R. (1992) Domestic Structure and Preventive War: Are Democracies More Pacific? *World Politics* 44 (2): 235–269.

Silverstone, S. (2007) *Preventive War and American Democracy*. New York: Routledge.

Skaperdas, S. (2002). Warlord Competition. *Journal of Peace Research* 39 (4): 435–446.

Skaperdas, S., and C. Syropoulos. (1996) Can the Shadow of the Future Harm Cooperation? *Journal of Economic Behavior and Organization* 29 (3): 355–372.

Slonim, R., and A. Roth. (1998) Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic. *Econometrica* 66 (3): 569–596.

Snidal, D. (1985) Coordination Versus Prisoners' Dilemma: Implications for International Cooperation and Regimes. *The American Political Science Review* 79 (4): 923–942.

Streich, P., and J. Levy. (2007) Time Horizons, Discounting, and Intertemporal choice. *Journal of Conflict Resolution* 51 (2): 199–226.

Tammen, R., J. Kugler, D. Lemke, M. Abdollohian, C. Alsharabati, B. Efird, A. Stam, and A. Organski. (2000) *Power Transitions: Strategies for the 21st Century*. New York: Chatham House Publishers.

Tingley, D., and B. Walter. (2011) The Effect of Repeated Play on Reputation Building An: Experimental Approach. *International Organization* 65 (2): 343–366.

Tingley, D., and W. Walter. (forthcoming) Can Cheap Talk Deter?. An Experimental Analysis. *Journal of Conflict Resolution*.

Tingley, D., and S. Wang. (2010) Belief Updating in Sequential Games of Two-Sided In-complete Information: An Experimental Study of a Crisis Bargaining Model. *Quarterly Journal of Political Science* 5 (3): 243–255.

Toft, M. D. (2007) Issue Indivisibility and Time Horizons as Rationalist Explanations for War. *Security Studies* 15 (1): 34–69.

Tomz, M. (2007) Domestic Audience Costs in International Relations: An Experimental Approach. *International Organization* 61 (4): 821–840.

Van Evera, S. (1985) Why Cooperation Failed in 1914. *World Politics* 38 (1): 80–117.

Walter, B. (1999) Designing Transitions From Civil War. *International Organization* 24 (1): 127–155.

Weg, E., A. Rapoport, and D. Felsenthal. (1990). Two-Person Bargaining Behavior in Fixed Discounting Factors Games with Infinite Horizon. *Games and Economic Behavior* 2 (1): 76–95.

Wilkenfeld, J., K. Young, V. Asal, and D. Quinn. (2003) Mediating International Crises: Cross-National and Experimental Perspectives. *Journal of Conflict Resolution* 47 (3): 279–301.

Zwick, R., and X. Chen. (1999) What Price Fairness? A Bargaining Study. *Management Science* 45 (6): 804–823.