Crisis Bargaining With Honor Considerations

Ahmer Tarar
Department of Political Science
Texas A&M University
email: ahmertarar@tamu.edu

May 23, 2016

Abstract

Some recent works in international relations argue that the field has, to its detriment, ignored Thucydides' emphasis of the role of honor in explaining war. I incorporate honor considerations into a rationalist model of crisis bargaining by supposing that a country pays an "honor cost" for accepting a small negotiated settlement to avoid war. The analysis identifies two mechanisms by which honor considerations can lead to costly war. First, the disputants' marginal honor costs may be jointly so high that the bargaining range is eliminated entirely and war occurs even under complete information. Second, even if the bargaining range is not eliminated, uncertainty about the other side's marginal honor cost can result in bargaining breakdown. I apply these results to the role of honor concerns in the outbreak of war between Athens and Melos during the Peloponnesian War, as well as recent negotiations over Iran's nuclear program.

1 Introduction

In the Melian Dialogue (*History of the Peloponnesian War*), the Greek historian Thucydides describes imperial Athens as offering Melos, in the words of the Athenian negotiator, "reasonable terms — alliance on a tribute-paying basis and liberty to enjoy your own property", to avoid war (Thucydides 1972, 406). The Melians propose (p.407) a counteroffer that Athens "allow us to be friends of yours and enemies to neither side, to make a treaty which shall be agreeable to both you and us, and so to leave our country." Negotiations fail and war breaks out, and one of the reasons the Melians give (p.403) for rejecting the Athenian proposal is that "we who are still free would show ourselves great cowards and weaklings if we failed to face everything that comes rather than submit to slavery."

The historian Donald Kagan (1995, 7-8) argues that Thucydides provided the most profound list of motives for why countries go to war: for reasons of "honor, fear, and interest." But he suggests that while "fear" and "interest" have been incorporated into modern theories of the causes of war, "honor" has been largely ignored, because modern analysts like to focus on "impersonal forces" (p.6). He goes on to write (p.8) that:

> The reader may be surprised by how small a role in the instances studied here, and, I believe, in many other cases, considerations of practical utility and material gain, and even ambition for power itself, play in bringing on wars and how often some aspect of honor is decisive.

Similarly, Lebow (2008, 131) laments the lack of attention paid to honor in contemporary social science and international relations theory and argues that "affronts to honor, and thus to self-esteem, have been at least as great a source of war as threats to material well-being

or security."[1]

Kagan and Lebow argue that concerns for honor have been prevalent through modern times, and not just in ancient Greece. For example, issues regarding honor seemed to play an important role in the recent negotiations on Iran's nuclear program. The *New York Times* reports that "As the negotiations sputtered forward, it became clear that to reach an agreement at all, Iran would have to be able to preserve a narrative of not backing down, not dismantling."[2] American negotiators "talk, in a wonderfully American way, about numbers and limits... [Iranian] officials talk almost entirely about preserving respect for their rights and Iran's sense of sovereignty... 'We are all about quantifiables: how many centrifuges can spin, how much plutonium can come out of the Arak reactor, how much uranium you can have on hand,' one senior American official... said... 'They are all about symbolism, about avoiding the optics of backing down,' the official said, even if it means engaging in expensive, inefficient nuclear enrichment activity that makes little economic or strategic sense."[3] Iranian foreign minister Mohammad Javad Zarif stated that "Our friends need to decide whether they want to be with Iran based on respect or whether they want to continue based on pressure... They have tested the other one; it is high time to test this one."[4]

In this paper, I present a game-theoretic analysis of how honor considerations can affect

[1] Referring to the related but more acquisitive notion of prestige, Markey (1999, 171) writes: "Today, however, mainstream international relations theory has dropped the 'prestige' variable from the political equation. The loss is palpable."

[2] "White Board, All-Nighters and Espresso" by David E. Sanger and Michael R. Gordon, *New York Times*, 4/4/2015, p.A1.

[3] "As Talks Drag On, US and Iran Find It Harder to Hear Each Other" by David E. Sanger, *New York Times*, 4/1/2015, p.A7.

[4] "Obama Told Negotiators to Disregard Deadline in Last Days of Iran Talks" by Michael R. Gordon and David E. Sanger, *New York Times*, 4/2/2015, p.A6.

crisis bargaining. In the rationalist bargaining model of war, the costliness of war means that there exists a range of negotiated settlements that both sides prefer to war, and the boundaries of this bargaining range are determined by material factors such as the military balance and the costs of war (Brito and Intriligator 1985; Bueno de Mesquita and Lalman 1992; Fearon 1995; Morrow 1989; Powell 1999). I modify the standard model so that a country's payoff for a negotiated settlement is not just its share of the disputed good in that settlement, but its share minus an "honor cost" that is increasing in the other side's share (i.e., decreasing in one's own share). This is meant to capture the idea that voluntarily accepting a small share carries a cost associated with one's honor, where honor can either be tied to an instrumental (material) loss due to a loss of reputation and its future negative consequences, or instead an inherent (perhaps psychological) cost which is not primarily linked to future reputation but instead to self-esteem. Kagan and Lebow have the latter in mind when referring to the importance of honor, rather than instrumental reputations for resolve that have been extensively studied in both the economics (Kreps and Wilson 1982; Milgrom and Roberts 1982) and international relations literatures (e.g., Huth 1997; Sechser 2010; Treisman 2004; Walter 2009; Weisiger and Yarhi-Milo 2015).[5] But the analysis in general allows for either possibility, where the reputational effect is captured in reduced form by not modeling any future interactions but simply assuming that accepting a small share,

[5]Markey (1999, 126; emphasis added) discusses how modern realists since Hans Morgenthau give at best an instrumental role to prestige, and Markey builds on classical realist philosophers to develop his conception of the "prestige motive" as "the individual or collective desire for public recognition of eminence *as an end in itself*." Similarly, Kagan (1997, 43; emphasis added) writes: "But nations, like individuals, uphold other conceptions of honor as well, and they also pursue honor in ways that are the product *not of calculation but of feeling*."

by signaling low resolve (e.g., a high cost of war), imposes a cost due to future reputational consequences.

I modify the ultimatum crisis bargaining model of Fearon (1995), in which country A makes a take-it-or-leave-it proposal to country B. When country B has the honor payoffs for accepting a negotiated settlement, then the minimum share it needs to avoid war (i.e., its reservation value) is higher than it otherwise would be, and hence the bargaining range is smaller than normal. Under complete information, if country B's marginal honor cost is not too high then country A makes an acceptable proposal and war is avoided. However, if the marginal honor cost exceeds a certain threshold, then honor considerations eliminate the bargaining range entirely and war occurs even under complete information. Empirically, it is unlikely that the marginal honor cost would be so high, but when both sides have the honor payoffs for accepting a negotiated settlement, then even moderate marginal honor costs can combine to eliminate the bargaining range. When honor considerations are involved, the boundaries of the bargaining range are determined not just by material factors such as the military balance and the costs of war (i.e., p, c_A, and c_B, in the usual crisis bargaining notation), but by the marginal honor costs as well, whose subjectivity and amorphousness make the interaction highly prone to informational problems.[6]

In this spirit, I then analyze the scenario where a bargaining range still exists but country A is uncertain of country B's marginal honor cost. For a broad set of parameter values, A's optimal proposal entails a risk of rejection and hence war. Thus, uncertainty about the other

[6]For example, O'Neill (1999, xii) writes: "Honor is an internal quality of the individual that can only be estimated by the rest of society. It involves, in part, the individual's desire to be seen as honorable, and this self-referential property is the key to proving one's honor and judging another person's honor."

side's marginal honor cost can lead to the outbreak of war, due to an informational problem and associated incentives to misrepresent one's private information (Fearon 1995).

Thus, the analysis identifies two mechanisms by which honor considerations can lead to costly war. First, honor considerations can be so large that they eliminate the bargaining range and lead to war even under complete information. This is unlikely to be the case in most empirical instances where honor considerations are involved, but even when a bargaining range still exists, uncertainty about the other side's marginal honor cost can lead to bargaining breakdown.

For example, it may have been that the Athenian proposal was within the bargaining range defined simply by the balance of military forces, which overwhelmingly favored Athens, and the costs of war, but was not within the true bargaining range that also incorporated the Melians' honor concerns. In fact, this analysis provides an alternative explanation for why strong states are often unsuccessful at peacefully compelling weak states to make policy concessions. Sechser (2010) identifies this empirical trend and provides an instrumental reputational explanation, using a two-period game-theoretic model to show that making large concessions now can signal that the weak state has a high cost of war and is thus willing to make large concessions in the next interaction as well. The alternative explanation provided here is that strong states may simply underestimate the honor concerns that smaller states have for making large concessions despite the military imbalance, and hence bargain simply on the basis of power considerations and without taking into account honor concerns. I will argue that a crucial reason for why the US and Iran were able to reach a negotiated settlement was because US negotiators, unlike their Athenian counterparts, recognized Iran's honor concerns for avoiding the appearance of backing down and took that into account while

bargaining.

Within the international relations literature, a number of recent works analyze the role of honor considerations (see especially Lebow 2008), and Dafoe, Renshon, and Huth (2014) review this and related literature. To my knowledge, the only game-theoretic treatment is O'Neill (1999). He provides a rich and nuanced analysis of honor and related considerations in international relations, but does not analyze the mechanisms by which honor considerations can lead to the outbreak of war in the rationalist bargaining model of war, which is my goal here.

The rest of the paper is organized as follows. In the next section, I present the game-theoretic analysis. I then discuss how the results apply to the Athens-Melos and US-Iran interactions. Next, I discuss how game theory can be useful for developing the precise causal mechanisms for how other non-rationalist factors can lead to the outbreak of war. Following that, I provide some additional applications of the model, other than honor considerations. Finally, I offer some concluding thoughts.

2 A Game-Theoretic Analysis

2.1 Only Country B Has Honor Considerations

I build on the following standard crisis bargaining setting. Two countries, A and B, have to divide a divisible good of value 1 between them, or go to war to decide who gets all of it. If war occurs, A wins with probability $p \in (0,1)$ and B wins with probability $1-p$. The costs of war are $c_A, c_B > 0$. Then, $EU_A(war) = (p)(1) + (1-p)(0) - c_A = p - c_A$ and $EU_B(war) = (p)(0) + (1-p)(1) - c_B = 1 - p - c_B$. Thus, $[p - c_A, p + c_B]$ is the bargaining range, i.e., the set of agreements that both sides prefer to war (Fearon 1995).

The bargaining protocol is a simple ultimatum proposal: A makes a take-it-or-leave-it proposal $(x, 1-x)$, where $x \in [0,1]$ is A's proposed share. B's two choices are to (i) accept this proposal, resulting in each side's payoff being its proposed share (i.e., assuming risk-neutrality), or (ii) go to war, in which case each side gets its expected payoff for war. The following is the standard result.

Proposition 1 *(Fearon 1995) This game has a unique subgame-perfect equilibrium (SPE) in which B accepts any proposal such that $x \leq p + c_B$, and A proposes $x = p + c_B$.*

Now suppose that B's payoff for accepting proposal $(x, 1-x)$ is $1 - x - h_B \cdot x$ rather than $1-x$, where $h_B \geq 0$ is a fixed constant that is B's marginal honor cost for reaching an agreement. That is, for reaching an agreement, B pays an honor cost $h_B \cdot x$ that is increasing in A's share, or alternatively decreasing in B's share: the more of the disputed good that B gets, the smaller its honor cost. With these modified payoffs, B accepts proposal $(x, 1-x)$ if $1 - x - h_B \cdot x \geq 1 - p - c_B$, which is equivalent to $x \leq \frac{p+c_B}{1+h_B} \in (0, p + c_B]$. The new bargaining range is thus $[p - c_A, \frac{p+c_B}{1+h_B}]$, whose width (in particular, the right boundary) is strictly decreasing in h_B. Note that $p - c_A \leq \frac{p+c_B}{1+h_B}$ (i.e., a bargaining range still exists) is equivalent to $h_B \leq \frac{c_A+c_B}{p-c_A} (> 0)$,[7] which leads to the following result.

Proposition 2 *This game has a (generically) unique subgame-perfect equilibrium (SPE) in which B accepts any proposal such that $x \leq \frac{p+c_B}{1+h_B}$, and A's proposal is as follows:*

(i) If $0 \leq h_B \leq \frac{c_A+c_B}{p-c_A}$, then A proposes $x = \frac{p+c_B}{1+h_B}$, which B accepts.

(ii) If $\frac{c_A+c_B}{p-c_A} < h_B$, then A makes any unacceptable proposal $x > \frac{p+c_B}{1+h_B}$ (this allows for non-uniqueness), and war results.

[7] I am assuming that $p - c_A > 0$ and $p + c_B < 1$, i.e., each side has a positive expected payoff for war and hence needs at least some of the disputed good to avoid war.

If the marginal honor cost h_B is not too high (case i), then a bargaining range still exists and war is avoided, with B getting more than it would without honor considerations. But if the marginal honor cost exceeds a certain threshold, then the bargaining range is eliminated and war occurs even under complete information.

Note by inspection that the critical threshold $\frac{c_A+c_B}{p-c_A}$ increases, and hence war is less likely (i.e., it is more likely that a bargaining range still exists), as (a) p decreases, (b) c_A increases, and (c) c_B increases. These are all intuitive comparative statics.

Empirically, it is unlikely that the marginal honor cost will be so high as to eliminate the bargaining range, but note that if the total cost of war $c_A + c_B$ is small, then h_B does not have to be very high for $\frac{c_A+c_B}{p-c_A} < h_B$ to be satisfied and hence war to occur. This suggests that when war is not expected to be very costly, honor concerns can easily eliminate the bargaining range.

2.2 Both Sides Have Honor Considerations

Now suppose that both sides face honor costs for reaching an agreement. That is, suppose that A's payoff if proposal $(x, 1-x)$ is accepted is $x - h_A(1-x)$ rather than x, where $h_A \geq 0$ is country A's marginal honor cost. For example, the main reason the Athenians gave (Thucydides 1972, 402-403) for why they required Melos to accept an unfavorable deal (tribute-paying alliance rather than neutrality) was that:

> ... if we were on friendly terms with you, our subjects would regard that as a sign of weakness in us, whereas your hatred is evidence of our power... by conquering you we shall increase not only the size but the security of our empire. We rule the sea and you are islanders, and weaker islanders too than the others; it is

therefore particularly important that you should not escape.

That is, the Athenians had a reputational motive to appear strong in order to deter rebellion, which in reduced form can be captured by a marginal honor cost: the more Melos gets in a deal, the greater Athens' reputational cost and future negative consequences.

Now country A prefers over war any agreement such that $x - h_A(1-x) \geq p - c_A$, which is equivalent to $x \geq \frac{p-c_A+h_A}{1+h_A} \in [p-c_A, 1)$. Now the bargaining range is $[\frac{p-c_A+h_A}{1+h_A}, \frac{p+c_B}{1+h_B}]$, whose width is strictly decreasing in both h_A and h_B.[8] This leads to the following result.

Proposition 3 *This game has a (generically) unique subgame-perfect equilibrium (SPE) in which B accepts any proposal such that $x \leq \frac{p+c_B}{1+h_B}$, and A's proposal is as follows:*

(i) If $\frac{p-c_A+h_A}{1+h_A} \leq \frac{p+c_B}{1+h_B}$, then A proposes $x = \frac{p+c_B}{1+h_B}$, which B accepts.

(ii) If $\frac{p-c_A+h_A}{1+h_A} > \frac{p+c_B}{1+h_B}$, then A makes any unacceptable proposal $x > \frac{p+c_B}{1+h_B}$, and war results.

The condition under which war occurs, $\frac{p-c_A+h_A}{1+h_A} > \frac{p+c_B}{1+h_B}$, is equivalent to $h_A[1-(p+c_B)] + h_B(p-c_A) + h_B \cdot h_A > c_A + c_B$. The left-hand-side of the latter inequality is strictly increasing in both h_A and h_B, meaning that if they are jointly sufficiently large, then a bargaining range does not exist and hence war occurs even under complete information. For example, it may have been that Athens' need to get a very favorable deal (or war) in order to maintain its instrumental reputation for toughness so as to deter colonial revolt, alongside Melos' inherent honor concern with accepting an unfavorable deal that it regarded as slavery, combined to eliminate the bargaining range and made war inevitable.

[8]Note that $\frac{\partial}{\partial h_A}[\frac{p-c_A+h_A}{1+h_A}] = \frac{1-(p-c_A)}{(1+h_A)^2} > 0$, i.e., A's reservation value, which is also the left boundary of the bargaining range, is strictly increasing in h_A. L'Hospital's Rule gives that $\lim_{h_A \to \infty}[\frac{p-c_A+h_A}{1+h_A}] = 1$.

Although it is unlikely in most empirical instances that honor concerns are so large as to eliminate the bargaining range, we can see from the latter inequality that, just as in the one-sided case, if the total cost of war $c_A + c_B$ is small, then h_A and h_B do not have to be very high at all to eliminate the bargaining range. Honor concerns can easily lead to war when the war is not expected to be very costly.

2.3 Uncertainty About Country B's Honor Considerations

Now return to the one-sided case where only country B has honor considerations (i.e., assume that $h_A = 0$), and suppose that A is uncertain about h_B. Suppose that h_B comes from a uniform distribution on the interval $[0, H_B]$, where $H_B > 0$. I also assume that $H_B \leq \frac{c_A + c_B}{p - c_A}$, so that a bargaining range exists even with type H_B, meaning that it also exists for every type $0 \leq h_B < H_B$. Thus, it is common knowledge that a bargaining range exists, but A is uncertain about h_B and hence the right boundary of the bargaining range.

In the first move of the game, "nature" chooses B's type from the distribution $U[0, H_B]$. B observes this move and hence knows its own marginal honor cost, but A does not observe that move and only knows the probability distribution when it makes its proposal. The following is the result.

Proposition 4 *This game has a unique perfect-Bayesian equilibrium (PBE), in which any type h_B of B accepts any proposal such that $x \leq \frac{p + c_B}{1 + h_B}$, and A's proposal is as follows:*

(i) If $0 < H_B \leq \sqrt{\frac{p + c_B}{p - c_A}} - 1$, then A makes the risk-free, limited proposal of $x = \frac{p + c_B}{1 + H_B}$, which all types of B accept.

(ii) If $\sqrt{\frac{p + c_B}{p - c_A}} - 1 < H_B$, then A makes the larger, risky proposal $x = \sqrt{(p - c_A)(p + c_B)} \in (\frac{p + c_B}{1 + H_B}, p + c_B)$, that not all types of B accept.

Under incomplete information, A faces a standard "risk-return tradeoff": the less it offers to B, the less likely the proposal is to be accepted, but the more A gains *if* it is accepted (Fearon 1995; Powell 1999). With a uniform distribution, if H_B is small (case i) and hence the distribution of types is narrow, then by making an even slightly risky (interior) proposal, the probability of rejection builds up too rapidly for the small gains (if the proposal is accepted) to be worthwhile, and hence A makes a safe proposal that all types of B accept. On the other hand, if H_B exceeds a certain threshold and hence the distribution of types is sufficiently wide (case ii), then A makes a risky proposal that not all types of B accept, and hence there is a risk of war.

Thus, private information provides a second mechanism by which honor considerations can lead to war. Also note that the incentive to misrepresent one's private information that is a crucial part of the informational explanation for costly conflict (Fearon 1995) applies here: because a larger h_B causes A to offer more to B (up to the point where h_B is so large that the bargaining range no longer exists), low-h_B types of B have incentives to bluff and claim that h_B is higher than it really is (i.e., that they are more concerned with honor than they really are), and this will make credible signaling difficult. For this reason, the Athenians had reasons to doubt the Melian claim that they would rather enter a (likely) disastrous war with Athens than accept the Athenian proposal of tribute-paying alliance.

3 Discussion

3.1 War Between Athens and Melos

The analysis provides two honor-based explanations that can account for the outbreak of war between Athens and Melos. It may have been that the honor considerations of the two

sides—an instrumental reputational one on Athens' part, and an inherent self-esteem one on Melos' part—combined to eliminate the bargaining range entirely and made war unavoidable. Alternatively, it may have been that a bargaining range still existed, but Athens underestimated Melos' honor considerations despite the latter's attempt to communicate them and hence made a too-harsh proposal that lay within the bargaining range determined purely by material power considerations, but outside of the true bargaining range that accommodated the Melians' honor concerns.

Many of the statements in the Melian Dialogue point to the informational explanation being more relevant, in that the Athenians repeatedly dismissed non-reputational honor concerns. For example, after the Melians appealed that their honor required them to reject the Athenian proposal, the Athenians responded (Thucydides 1972, 403-404, 406):

> This is no fair fight, with honor on one side and shame on the other. It is rather a question of saving your lives and not resisting those who are far too strong for you... Do not be led astray by a false sense of honor — a thing which often brings men to ruin when they are faced with an obvious danger that somehow affects their pride. For in many cases men have still been able to see the dangers ahead of them, but this thing called dishonor, this word, by its own force of seduction, has drawn them into a state where they have surrendered to an idea, while in fact they have fallen voluntarily into irrevocable disaster, in dishonor that is all the more dishonorable because it has come to them from their own folly rather than their misfortune.

Similarly, at one point the Melians justify their hope that Athens' enemy Sparta will come to Melos' aid by stating (p.404) that the Spartans "... are bound, if for no other reason, then

for honor's sake, and because we are their kinsmen, to come to our help." The Athenians reply (p.405) that "... with regard to your views about Sparta and your confidence that she, out of a sense of honor, will come to your aid, we must say that we congratulate you on your simplicity but do not envy you your folly... of all people we know the Spartans are most conspicuous for believing that what they like doing is honorable and what suits their interest is just."

In-so-far as these statements can be interpreted as the Athenians downplaying the relevance of non-instrumental, inherent honor, they point to the informational explanation as being more relevant in this case: the Athenians simply do not take such honor concerns seriously and hence are bargaining purely on the basis of power considerations. However, another interpretation is that the Athenians (or Thucydides, through their voice) are recognizing the *empirical* reality of such honor considerations, but are making an argument about the *wisdom* of allowing honor to cloud one's judgement. Under this interpretation, the complete-information explanation is more relevant, with the Athenians recognizing that the joint honor concerns are combining to eliminate the bargaining range, and are trying to talk the Melians down from their high honor concern in order to get them to accept Athens' terms.

3.2 Agreement Between the US and Iran

In contrast, in the recent negotiations over Iran's nuclear program, the analysis suggests that a crucial reason for why an agreement was able to be reached was because US negotiators, unlike their Athenian counterparts, recognized Iran's real honor concerns for avoiding the appearance of having backed down, and negotiated an agreement that accommodated those concerns instead of bargaining purely on the basis of power considerations.

For example, the *New York Times* reports that a crucial hurdle to reaching an agreement was Iran's demand that it be allowed to keep 1,000 centrifuges at the Fordo nuclear facility. The US relented after Iran agreed that no uranium enrichment would occur at the facility, and "[Obama] administration officials were struck by the fact that Iran was willing to waste 1,000 centrifuges, essentially spinning uselessly, to preserve national pride."[9] According to the analysis presented here, the Obama administration's willingness to accommodate such honor concerns was crucial to an agreement being reached.[10]

3.3 Using Game Theory to Elucidate the Causal Mechanisms for Non-rationalist Explanations for War

The approach adopted in this paper, namely incorporating a non-rationalist, perhaps psychological, factor—non-instrumental, self-esteem based honor—into the utility function of an actor in a game-theoretic model, may be a fruitful way of theorizing about the causal mechanisms for some other non-rationalist factors as well. Psychological factors undoubtedly play an important role in many decisions regarding war and peace, and one of the the strengths of game-theoretic modeling is in identifying precise causal mechanisms. In this paper, I have identified two causal mechanisms for how honor concerns can lead to war, without trying to explain the origins of honor concerns and instead simply taking them as given.[11]

[9] "White Board, All-Nighters and Espresso" by David E. Sanger and Michael R. Gordon, *New York Times*, 4/4/2015, p.A1.

[10] "The provision, Obama administration officials assert, carries no serious risk for the United States but will enable the Iranians to save face." In "Outline of Nuclear Agreement Sounds Different From Each Side" by Michael R. Gordon, *New York Times*, 4/5/2015, p.A8.

[11] As Markey (1999) points out for the related notion of prestige, Hobbes believes that the desire for prestige is rooted in human nature, whereas Rousseau believes that it is the result of socialization.

In a similar vein, Fearon (1995) identified issue indivisibility as a rationalist explanation for war, but downplayed its importance because territory is almost always physically divisible and even with genuinely indivisible goods, issue linkage and side payments can often "induce" divisibility. However, others have argued that Fearon dismissed issue indivisibility too quickly, because non-rationalist factors can account for how actors often empirically come to view territory as indivisible (e.g., Goddard 2006; Hassner 2007; Toft 2005). Firmly within the rationalist tradition, Powell (2006) further developed the causal mechanism by showing that when an issue viewed as indivisible—perhaps due to non-rationalist factors—leads to war, it is ultimately due to a commitment problem, namely the inability of the disputants to commit to complying with the result of a non-costly lottery for allocating the indivisible good. Similarly, I have argued that when another non-rationalist factor—non-instrumental, self-esteem based honor—leads to war, it is either through eliminating the bargaining range or due to an informational problem.

Thus, game-theoretic modeling is a promising way of identifying precise causal mechanisms for how even non-rationalist, perhaps psychological, factors can lead to war, especially those factors that can be interpreted as applying to the actors' payoffs (e.g., honor concerns) or the nature of the strategic interaction (e.g., indivisible goods), as opposed to the decision-making process itself (e.g., group think).

3.4 Additional Applications of the Model

The model analyzed in this paper is ultimately a model of crisis bargaining in which negotiating is costly (and not just war, as in most bargaining models of war), and there are other possible applications of the model, besides honor considerations. For example, this could be interpreted as a model in which one or both leaders have publicly portrayed the disputed

good as being indivisible, and hence dividing it incurs domestic costs for the leader(s). Under this interpretation, the analysis suggests that even if the disputed good is not literally indivisible, portraying it as such can lead to war if the marginal cost (either one-sided or jointly) for then dividing it is large enough as to eliminate the bargaining range, or if there is uncertainty about the other side's marginal cost.[12]

Alternatively, the model can be interpreted as a crisis bargaining model in which a hawkish party, typically an opposition party but possibly the leader's own party, imposes costs on a dovish or moderate leader for negotiating with an adversary. Interpreted this way, the analysis suggests that a hawkish opposition can lead to the country getting a better deal in crisis bargaining, which is consistent with the "Schelling conjecture" in the two-level game literature on legislative ratification of treaties (Milner 1997; Putnam 1988; Schelling 1960). But this is the case only within limits, as too-large marginal negotiating costs eliminate the bargaining range altogether, and uncertainty about even moderate marginal costs can lead to bargaining breakdown. In addition, if only dovish or moderate leaders pay such negotiating costs (only a Nixon can go to China), then the analysis suggests somewhat counter-intuitively that hawkish leaders are more able to reach war-avoiding agreements with adversaries.[13]

[12]Tarar and Leventoğlu (2009) analyze a model in which a leader endogenously chooses how much of the disputed good to publicly commit to obtaining in negotiations, and pays a cost only if the leader gets less than what she committed to obtaining. That captures the audience cost idea that domestic actors dislike inconsistency between a leader's statements and actions, and hence the leader can *generate* potential negotiating costs by making public statements (Fearon 1994). In this paper, it is assumed that negotiating costs simply exist, regardless of the leader's attempt to manipulate them.

[13]Schultz (2005) has a similar result. On the other hand, hawkish leaders have lower costs of war, which leads to a higher probability of war in risk-return equilibria (for example, see the comparative statics section in the appendix).

4 Conclusion

A number of authors have argued that the international relations field has, to its detriment, ignored the importance of honor considerations in decisions leading to war and peace. In this paper, I have incorporated honor considerations into a rationalist crisis bargaining model. The analysis identifies two mechanisms by which honor concerns can lead to war. First, sufficiently large marginal honor costs can eliminate the bargaining range entirely and lead to war even under complete information. Empirically, it is unlikely that honor concerns will be so large, but when the total (expected) costs of war are low, then even moderate marginal honor costs can eliminate the bargaining range. Second, uncertainty about the other side's marginal honor cost can lead to bargaining breakdown and hence war. One or both of these mechanisms may have contributed to the outbreak of war between Athens and Melos during the Peloponnesian War. In contrast, I have argued that a crucial reason for why an agreement was reached between the US and Iran was because US negotiators, unlike their Athenian counterparts, took into account Iran's real honor concerns for avoiding the appearance of having backed down, and did not bargain purely on the basis of power considerations.

I have also argued that game theory can help identify the precise causal mechanisms for how other non-rationalist factors can lead to war. Finally, the model can alternatively be interpreted as a model of crisis bargaining in which one or both leaders have publicly portrayed a divisible disputed good as being indivisible, or as one in which one or both leaders face hawkish oppositions that oppose and punish negotiations. The model thereby sheds some insight into these two strategic settings as well.

5 Appendix

Proposition 1 is a standard result, and Propositions 2 and 3 are essentially derived in the main text. Hence, I just provide a proof of Proposition 4, and then provide comparative statics on it.

5.1 Proof of Proposition 4

Any proposal $x \in [0, \frac{p+c_B}{1+H_B}]$ is accepted for sure, resulting in A's payoff being x. Within this range, A's uniquely optimal proposal is thus $x = \frac{p+c_B}{1+H_B}$. Also note that A's payoff for any $x \geq p+c_B$ is constant at $p-c_A$, since any such proposal is rejected with probability 1. Thus, we just need to maximize over the interval $x \in [\frac{p+c_B}{1+H_B}, p+c_B]$, i.e., the optimal value of x occurs in this interval.

For any $x \in [\frac{p+c_B}{1+H_B}, p+c_B]$, B accepts the proposal if $x \leq \frac{p+c_B}{1+h_B}$, which is equivalent to $h_B \leq \frac{p+c_B-x}{x}$, which occurs with probability $F(\frac{p+c_B-x}{x})$, where $F(\cdot)$ is the cumulative distribution function for h_B. If the proposal is accepted, A's payoff is x. B rejects the proposal if $x > \frac{p+c_B}{1+h_B}$, which is equivalent to $h_B > \frac{p+c_B-x}{x}$, which occurs with probability $1-F(\frac{p+c_B-x}{x})$. If the proposal is rejected, A's payoff is $p-c_A$. Thus, A's expected utility for any $x \in [\frac{p+c_B}{1+H_B}, p+c_B]$ is $EU_A(x) = F(\frac{p+c_B-x}{x})(x)+[1-F(\frac{p+c_B-x}{x})](p-c_A) = F(\frac{p+c_B-x}{x})(x-p+c_A)+(p-c_A)$. Differentiating, $EU'_A(x) = F(\frac{p+c_B-x}{x}) - f(\frac{p+c_B-x}{x})[\frac{p+c_B}{x^2}][x-(p-c_A)]$.

Substituting in the uniform density and distribution functions $f(\frac{p+c_B-x}{x}) = \frac{1}{H_B}$ and $F(\frac{p+c_B-x}{x}) = \frac{p+c_B-x}{xH_B}$ gives $EU'_A(x) = \frac{(p-c_A)(p+c_B)-x^2}{x^2H_B}$, and setting this equal to 0 gives $x^* = \sqrt{(p-c_A)(p+c_B)} \in (p-c_A, p+c_B)$ (the negative root is obviously not relevant). Also note that $EU''_A(x) = \frac{-2xH_B(p-c_A)(p+c_B)}{x^4H_B^2}$, which is negative if and only if $x > 0$. Thus, $EU''_A(x) < 0$ over the entire interval $x \in [\frac{p+c_B}{1+H_B}, p+c_B]$, which means that $x^* = \sqrt{(p-c_A)(p+c_B)}$

maximizes rather than minimizes $EU_A(x)$. Finally, we need to determine the condition under which $x^* = \sqrt{(p-c_A)(p+c_B)} > \frac{p+c_B}{1+H_B}$, i.e., x^* is risky. This is equivalent to $H_B > \sqrt{\frac{p+c_B}{p-c_A}} - 1$ (> 0). On the other hand, if $H_B \leq \sqrt{\frac{p+c_B}{p-c_A}} - 1$, then $x^* = \sqrt{(p-c_A)(p+c_B)} \leq \frac{p+c_B}{1+H_B}$, meaning that A chooses the safe proposal of $x^{**} = \frac{p+c_B}{1+H_B}$. Q.E.D.

5.2 Comparative Statics

Note by inspection that $x^* = \sqrt{(p-c_A)(p+c_B)}$ is strictly increasing in p and c_B, and strictly decreasing in c_A, which are all intuitive results.

Note that $Pr(war) = 1 - F(\frac{p+c_B-x}{x})$, and when we substitute in the uniform distribution function and $x^* = \sqrt{(p-c_A)(p+c_B)}$, then we get $Pr(war) = \frac{(H_B+1)\sqrt{(p-c_A)(p+c_B)}-p-c_B}{H_B\sqrt{(p-c_A)(p+c_B)}}$.

Differentiation gives $\frac{\partial}{\partial c_A}Pr(war) = \frac{-(p+c_B)}{2H_B(p-c_A)\sqrt{(p-c_A)(p+c_B)}} < 0$, and $\frac{\partial}{\partial c_B}Pr(war) = \frac{-1}{2H_B\sqrt{(p-c_A)(p+c_B)}} < 0$. Hence, the equilibrium probability of war is strictly decreasing in c_A and c_B, which are standard results for risk-return equilibria.

Note that $\frac{\partial}{\partial H_B}Pr(war) = \frac{p+c_B-\sqrt{(p-c_A)(p+c_B)}}{H_B^2\sqrt{(p-c_A)(p+c_B)}} > 0$, and hence the equilibrium probability of war is strictly increasing in the variance of the distribution of types, which is consistent with Reed (2003) and Wittman (2009).

Finally, note that $\frac{\partial}{\partial p}Pr(war) = \frac{c_A+c_B}{2H_B(p-c_A)\sqrt{(p-c_A)(p+c_B)}} > 0$, and hence the equilibrium probability of war is strictly increasing in p. Thus, the "neutrality" result that often holds in risk-return equilibria, whereby an increase in p causes A to become more aggressive but B to become more conciliatory in a way that the probability of war remains constant (e.g., Cetinyan 2002; Kydd 2010; Kydd and Straus 2013; Wittman 1979), does not hold. Benson, Meirowitz, and Ramsay (2015) provide a detailed analysis of the conditions under which the neutrality result holds, and it suffices here to note that it does not hold in this model.

References

[1] Benson, Brett, Adam Meirowitz, and Kristopher Ramsay. 2015. "The Balance of Power and the Risk of War in Crisis Bargaining." Unpublished manuscript, Vanderbilt University and Princeton University.

[2] Brito, Dagobert L., and Michael D. Intriligator. 1985. "Conflict, War, and Redistribution." *American Political Science Review* 79(4):943-57.

[3] Bueno de Mesquita, Bruce, and David Lalman. 1992. *War and Reason*. New Haven: Yale University Press.

[4] Cetinyan, Rupen. 2002. "Ethnic Bargaining in the Shadow of Third-Party Intervention." *International Organization* 56(3):645-677.

[5] Dafoe, Allan, Jonathan Renshon, and Paul Huth. 2014. "Reputation and Status as Motives for War." *Annual Review of Political Science* 17:371-93.

[6] Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88(September):577-592.

[7] Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379-414.

[8] Goddard, Stacie. 2006. "Uncommon Ground: Indivisible Territory and the Politics of Legitimacy." *International Organization* 60(1):35-68.

[9] Hassner, Ron. 2007. "The Path to Intractability: Time and the Entrenchment of Territorial Disputes." *International Security* 31(3):107-38.

[10] Huth, Paul. 1997. "Reputations and Deterrence: A Theoretical and Empirical Assessment." *Security Studies* 7(1):72-99.

[11] Kagan, Donald. 1995. *On the Origins of War and the Preservation of Peace*. New York: Anchor Books.

[12] Kagan, Donald. 1997. "Our Interests and Our Honor." *Commentary* (April):42-45.

[13] Kreps, David, and Robert Wilson. 1982. "Reputation and Imperfect Information." *Journal of Economic Theory* 27:253-279.

[14] Kydd, Andrew H. 2010. "Rationalist Approaches to Conflict Prevention and Resolution." *Annual Review of Political Science* 13:101-121.

[15] Kydd, Andrew H., and Scott Straus. 2013. "The Road to Hell? Third-Party Intervention to Prevent Atrocities." *American Journal of Political Science* 57(3):673-684.

[16] Lebow, Richard N. 2008. *A Cultural Theory of International Relations*. New York: Cambridge University Press.

[17] Markey, Daniel. 1999. "Prestige and the Origins of War." *Security Studies* 8(4):126-73.

[18] Milgrom, Paul, and John Roberts. 1982. "Predation, Reputation, and Entry Deterrence." *Journal of Economic Theory* 27:280-312.

[19] Milner, Helen. 1997. *Interests, Institutions, and Information*. Princeton University Press.

[20] Morrow, James D. 1989. "Capabilities, Uncertainty, and Resolve: A Limited Information Model of Crisis Bargaining." *American Journal of Political Science* 33(November):941-972.

[21] O'Neill, Barry. 1999. *Honor, Symbols, and War*. University of Michigan Press.

[22] Powell, Robert. 1999. *In the Shadow of Power: States and Strategies in International Politics*. Princeton: Princeton University Press.

[23] Powell, Robert. 2006. "War as a Commitment Problem." *International Organization* 60:169-203.

[24] Putnam, Robert D. 1988. "Diplomacy and Domestic Politics: The Logic of Two-level Games." *International Organization* 42(Summer):427-460.

[25] Reed, William. 2003. "Information, Power, and War." *American Political Science Review* 97(4):633-641.

[26] Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge: Harvard University Press.

[27] Schultz, Kenneth A. 2005. "The Politics of Risking Peace: Do Hawks or Doves Deliver the Olive Branch?" *International Organization* 59(1):1-38.

[28] Sechser, Todd S. 2010. "Goliath's Curse: Coercive Threats and Asymmetric Power." *International Organization* 64(4):627-660.

[29] Tarar, Ahmer, and Bahar Leventoğlu. 2009. "Public Commitment in Crisis Bargaining." *International Studies Quarterly* 53(3):817-839.

[30] Thucydides. 1972. *History of the Peloponnesian War*. Translated by Rex Warner, edited by M. I. Finley. Penguin Classics.

[31] Toft, Monica. 2005. *The Geography of Ethnic Violence: Identity, Interests, and the Indivisibility of Territory*. Princeton, NJ: Princeton University Press.

[32] Treisman, Daniel. 2004. "Rational Appeasement." *International Organization* 58(2):345-73.

[33] Walter, Barbara F. 2009. *Reputation and Civil War: Why Separatist Conflicts Are So Violent*. New York: Cambridge University Press.

[34] Weisiger, Alex, and Keren Yarhi-Milo. 2015. "Revisiting Reputation: How Past Actions Matter in International Politics." *International Organization* 69(2):473-495.

[35] Wittman, Donald. 1979. "How a War Ends: A Rational Model Approach." *Journal of Conflict Resolution* 23(December):743-63.

[36] Wittman, Donald. 2009. "Bargaining in the Shadow of War: When is a Peaceful Resolution Most Likely?" *American Journal of Political Science* 53(3):588-602.