

LETTER

## What the Enemy Knows: Common Knowledge and the Rationality of War

Thomas Chadeaux

Department of Political Science, Trinity College Dublin. Email: thomas.chadeaux@tcd.ie

(Received 12 April 2016; revised 22 August 2016; accepted 24 April 2018; First published online 31 October 2018)

The onset of costly interstate wars, despite the existence of more efficient negotiated solutions, is a longstanding puzzle in international conflict research. Much of the literature points to private information as a central cause for these bargaining failures. The argument is typically that one or both states misestimate resolve, preferences, costs or the distribution of capabilities, and that this asymmetric information can lead both to expect a positive utility for war.

Often, however, the uncertainty is not about capabilities or resolve, but rather about uncertainty itself: how much does the enemy know, how much do they know about what I know, about what I know that they know, and so on. In other words, both parties are often well aware of what the other has or wants (first-order uncertainty), but not of what the other knows (second-order uncertainty). In the lead-up to the 2003 invasion of Iraq, for example, much of the uncertainty was about what the other knew itself. Saddam Hussein was not sure whether the United States knew about the absence of weapons of mass destruction, and he also worried about what Iran knew about Iraq's capabilities. This uncertainty about uncertainty itself contributed to misunderstandings, delays in complying and ultimately war.

Unfortunately, existing models are ill equipped to address this type of situation. The asymmetric information included in crisis bargaining models is typically limited to payoff-relevant events fundamentals such as capabilities, costs or preferences. What players do or do not know, however, is almost always assumed to be common knowledge.<sup>1</sup> States may not know what the other has, but they know whether she knows. There is, in other words, no uncertainty about the presence of uncertainty.

In the real world, however, diplomats constantly report about their foreign counterparts' thoughts and knowledge, and double agents feed the enemy potentially erroneous information about what is known and believed, all the while guessing whether their opponent's spies succeeded in acquiring secret information. The success of signals also depends on the other's understanding, as well as my understanding of their understanding, and so on. What the other knows, in other words, is itself private and potentially strategic information.

I show here that bargaining may in fact break down into war simply because of uncertainty about uncertainty itself, even if both states are mutually aware of each other's capabilities, costs and resolve. The intuition for this result is simple: suppose that *A* thinks *B* might be mistaken about the distribution of power (even if she is not). For example, *A* might think that *B* thinks *A* has developed nuclear capabilities, that is, that the balance of power is less favourable to *B* than it actually is. If it is sufficiently probable that *B* is mistaken, then *A* has an incentive to offer her a small share of the pie. But because there is a chance that *B* actually knows the truth, *B* might in fact prefer war to that offer.

---

<sup>1</sup>We say that a proposition is *common knowledge* if every player knows it, every player knows that every player knows it, and so on ad infinitum (Aumann 1976). A player's *private information* is any information that he has that is not common knowledge among all players in the game.

This result has important implications. If war is caused by incomplete information about fundamental attributes such as capabilities, resolve or costs, then transparency or credible signals should solve the problem of inefficient bargaining outcomes. But if complete information about the distribution of information is also necessary, then the conditions for the absence of conflict are more stringent than previously thought. In particular, it implies that states must not only gather and convey first-order information—about their opponents’ capabilities and resolve—but also second- or higher-order information: information about what the opponent knows about her, and possibly what he knows she knows about her, and so on. They also must be able to convey that information about what they know or do not know, each time with a chance of failure.

The result is also relevant for the empirical literature. Existing empirical evidence of the role of uncertainty in the onset of war relies on estimates of first-order uncertainty—for example, using secret mobilization—and their correlations with the onset of conflict.<sup>2</sup> However, the present article suggests that these inferences may be incorrect if information about what the other knows is not taken into account.

In sum, this note shows the importance of higher-order uncertainty for the onset of war. The *possibility* of uncertainty alone may lead two rational and unitary actors to fight, even if they know each other’s capabilities and resolve. The article proceeds in three steps. First, the role of information and common knowledge in the bargaining literature is discussed. Secondly, a simple model is introduced in which war occurs with positive probability in all equilibria simply because of uncertainty about uncertainty itself. The logic of the argument is illustrated using the case of the lead-up to the 2003 Iraq War. Having established the importance of higher-order information, I finally address some of the challenges associated with reaching common knowledge.

### Uncertain Uncertainty

The role of information in interstate relations has been at the core of the research on conflict over the past thirty years. The central explanation for the onset of war between rational actors is the idea that at least one player has incomplete information about some of their opponent’s attributes such as their capabilities, resolve or costs for war (Blainey 1988; Fearon 1995; Jervis 1976; Powell 1996; Reiter 2003).<sup>3</sup> Private information and incentives to misrepresent them, in turn, lead to misperceptions or miscalculations about the distribution of power or resolve. Negotiators may thus be optimistic about the expected outcome of a war, with the result that both believe they have a higher expected value for war than for peace (Fey and Ramsay 2007; Slantchev and Tarar 2011).<sup>4</sup>

The type of uncertainty that is typically modelled, however, relates only to payoffs. The games include some uncertainty space corresponding to payoff-relevant attributes, together with the information players have about these attributes. Blainey’s often-cited argument, for example, is that ‘wars usually begin when two nations disagree on their relative strength’ (Blainey 1988, 246). Similarly, Fearon (1995, 18) focuses on ‘disagreements about relative power and uncertainty about a

<sup>2</sup>Some of the proxies for uncertainty rely on the military balance (Bas and Schub 2016a; Reed 2003; Slantchev 2004), while others are based on private capabilities such as secret alliances (Bas and Schub 2016b) or secret mobilization (Lai 2004).

<sup>3</sup>Of course, not all of the bargaining literature is related to incomplete information – see, for example, rationalist explanations of war related to commitment problems (Chadeaux 2011; Chadeaux 2012; Powell 2004b) or indivisibilities (Toft 2006). Gartzke (1999), however, argues that the three rationalist explanations presented in Fearon (1995) – private information, commitment problems and indivisibilities – are all instances of inefficiencies caused by one form or another of incomplete information.

<sup>4</sup>Asymmetric information may also cause wars to last longer than they would otherwise (Filson and Werner 2002; Powell 2004a; Slantchev 2003; Smith and Stam 2004; Wagner 2000; Wittman 1979). See also Schub (2015) for an empirical test of the relationship between uncertainty and the onset of war.

potential opponent's willingness to fight'.<sup>5</sup> In these models players may be unsure of each other's attributes such as the distribution of capabilities or interest, or the cost of war, but they know whether each one is informed of them.<sup>6</sup> In other words, the type of uncertainty that is discussed is of the form 'A does not know X, and B knows that A does not know X'. The fact that information is asymmetric is common knowledge, so states only face what may be called first-order uncertainty—uncertainty about fundamentals, but not about the information partition itself.<sup>7</sup>

In the real world, however, states rarely know for sure how much the other knows. Intelligence failures and the challenges associated with credibly conveying private information imply that states must almost always guess what the other has learned or inferred. For example, has the adversary discovered my secret nuclear programme (or the absence thereof)? Are enemy spies leaking information? Is my phone tapped? And what is made of that information? Did they believe it? And did they think that I thought they believed it?

Can bargaining remain efficient if everything is known except whether the other knows? We know from economics that uncertainty about what others know matters.<sup>8</sup> Yet of interest in the literature has been the opposite scenario: one in which information is *incomplete* and trade never happens because *common knowledge* allows the actors to infer from an offer that the other must know something they do not, or else they would not want to trade (Milgrom and Stokey 1982). Through a potentially infinite regress, they infer that the expected utility of the deal cannot be larger than the one of no trade, and hence trade never occurs. Analogous arguments have been discussed with respect to conflict (Fey and Ramsay 2007; Slantchev and Tarar 2011). However, the argument in this article is the converse: even with complete information but without common knowledge, war may occur in equilibrium. I now show this using a simple model with only two types and one period. The more general case with an infinite number of types and repeated interactions is discussed below.

## Model

Two countries, *A* and *B*, negotiate over the partition of a territory of size normalized to one. There are two types of *A*, denoted by  $\phi \in \{w, s\}$ : a strong *A*, denoted  $A^s$ , would win a conflict against *B* with probability  $p^s$ , whereas a weak *A* ( $A^w$ ) would win with probability  $p^w$ . *A*'s type is determined by an initial move by Nature.

Both players know the distribution of power, but *A* is unsure whether *B* knows it. To represent this uncertainty, there must therefore be states of the world reached with positive probability in which *B* knows the distribution of power, and others in which *B* does not. I represent these by a

<sup>5</sup>Further, Fearon also writes: 'state leaders might have private information about militarily relevant factors—military capabilities, strategy, and tactics; the population's willingness to prosecute a long war; or third-state intentions. If a state has superior (and so private) information about any such factor, then its estimate of the probable course of battle may differ from that of an adversary'.

<sup>6</sup>Other contributions in this field adopt the same approach: Morrow models a situation in which 'each side is uncertain about the other's resolve' (1989, 942); Powell assumes that 'each player has private information about its cost of imposing a settlement' (1996, 261); in Fearon, 'the defender is privately informed of its value for the international interest and then chooses a signal  $m$  [...]' (1997, 70). In Schultz, 'each state is uncertain about the other's expected costs from war' (1998, 834). In Gartzke, 'the other country [...] has private information about its costs [of fighting]' (1999, 578). Wagner writes that 'in the case of war what prevents immediate agreement is not private information about the personal preferences of political leaders, but conflicting expectations about the relative performance of military forces and/or about the behavior of other actors such as potential allies or political actors within states' (2000, 472). Slantchev assumes that 'player 1 is uncertain about the distribution of power' (2003, 624), and Smith and Stam that 'although actors have different beliefs, what these beliefs are is common knowledge to all actors. That is, although nation *A* has different beliefs than nation *B*, nation *A* knows what nation *B* believes and vice versa' (2004, 787).

<sup>7</sup>A recent notable exception is Acharya and Grillo (2015), but their focus is on the absence of common knowledge of *rationality* rather than on higher-order information.

<sup>8</sup>Rubinstein (1989), for example, demonstrates that even Nash equilibria that survive the usual refinements may not be robust to a small lack of common knowledge.

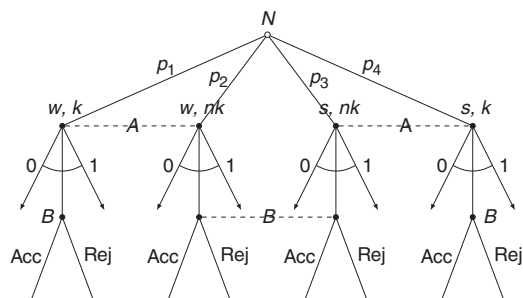


Figure 1. The common knowledge game: A does not know whether B knows A's type.

move by Nature, such that instead of having two possible types corresponding to two distributions of power, Nature has four possible moves, corresponding to the combination of A's two possible types ( $w$  and  $s$ ) and B's two possible states of knowledge: knows ( $k$ ) and not knows ( $nk$ )—see Figure 1.<sup>9</sup> In state  $\{w, k\}$ , both players know that A is weak, but A does not know whether B knows it. In state  $\{w, nk\}$ , A knows that she is weak, but B does not; again, A does not know whether B is informed. In state  $\{s, nk\}$ , A knows she is strong but B does not, and A again does not know whether B knows. Finally, in state  $\{s, k\}$ , both are informed that A is strong, but A again does not know whether B knows it. Assuming that all states have positive prior probability (that is,  $p_1, p_2, p_3, p_4 > 0$ ), only the entire state space is common knowledge.

In short, A always knows her own type, but never knows whether B observes A's type. Parentheses are used to denote a player's information partition. For example,  $(\{x\} \vee \{y\})$  denotes an information partition in which a player knows that she is at either  $x$  or  $y$ , but not which of the two. With probability  $p_1$ , for example, the state is  $w, k$ , in which case B observes  $(\{w, k\})$  and hence knows A's type. In this state, however, A only observes  $(\{w, k\} \vee \{w, nk\})$ , and hence assigns positive probability to state  $\{w, nk\}$  being the true state—a state in which B would not know A's type. Of particular interest here will be the states in which both players mutually know each other's capabilities, and yet end up fighting. They fight because, despite this mutual knowledge, the game is embedded in a larger structure in which capabilities may not always be known, and the mere *possibility* of asymmetric information—not asymmetric information itself—is sufficient to cause war.

Following Nature's move, I assume for simplicity a take-it-or-leave-it bargaining protocol in which A makes an offer  $x \in [0, 1]$ , where  $x$  denotes A's proposed share of the territory (and hence  $1 - x$  denotes B's share). B observes the offer, which she either accepts or rejects. If B accepts, then players receive their respective share of the pie and the game ends. If B rejects, however, then war follows, in which case A wins the entire territory with probability  $p^\phi$  (and hence B wins with probability  $1 - p^\phi$ ), and both players incur cost  $c \in (0, 1)$ . I assume for simplicity of exposition that both players are risk neutral (that is,  $u_i(x) = x$ ).

The intuition for the game's equilibrium is simple. Suppose that A observes  $(\{w, k\} \vee \{w, nk\})$ . This means that A is weak, but also that A is not sure whether or not B knows that she is weak, because A assigns positive probability to  $\{w, nk\}$  being the actual state. Indeed, if  $\{w, nk\}$  were the true state, B would observe  $(\{w, nk\} \vee \{s, nk\})$ , and hence would assign positive probability to the possibility that A is strong. In other words, upon observing  $(\{w, k\} \vee \{w, nk\})$ , A assigns positive probability to B thinking she is strong. But in that case (that is, if B thinks A might be strong), then B might be willing to accept a distribution of the pie that reflects this. In turn, this means that A could make an offer  $x$  corresponding to a strong type, which B would accept given her beliefs.

<sup>9</sup>We follow Harsanyi (2004) and model the absence of common knowledge using a game of incomplete information. See also Fudenberg and Tirole (1991, 556–7).

In equilibrium, therefore, *A* makes an offer that grants herself a large portion of the pie, with the hope that *B* does not actually observe the true distribution of power. Because *B* may, however, observe it (since with probability  $p_1$  the state is  $\{w, k\}$ ), war occurs with positive probability. For certain combinations of the parameters, I even show that war occurs with positive probability in all perfect Bayesian equilibria of the game.

Whether this can be an equilibrium strategy depends, of course, on the probability of  $\{w, k\}$  being the true state of the world given that *A* observes ( $\{w, k\} \vee \{w, nk\}$ ), as well as on  $\{s, k\}$  being the true state of the world when *B* observes ( $\{w, nk\} \vee \{s, nk\}$ ). In particular, it is crucial that the probability of being at  $\{w, k\}$  be sufficiently low when *A* observes ( $\{w, k\} \vee \{w, nk\}$ ), so that *A* would want to take the risk to bluff with the hope of not facing a knowledgeable *B*. It is also necessary that  $\{s, nk\}$  be sufficiently probable for *B* to accept upon observing ( $\{w, nk\} \vee \{s, nk\}$ ), or else she would rather take the risk to reject. I now present the logic of the equilibrium in more detail.

### Separating equilibrium

Consider first the possibility of an equilibrium in which each type of *A* makes a different offer. Without a loss of generality, assume that *A* offers  $x^L$  upon observing ( $\{w, k\} \vee \{w, nk\}$ ), but  $x^H$  upon observing ( $\{s, nk\} \vee \{s, k\}$ ), where  $x^L$  ( $x^H$ ) denotes a low (high) offer. In this case, *B* always infers from *A*'s offer which type she is facing, and war never occurs. But does such an equilibrium exist? A weak *A* may be tempted to deviate and offer  $x^H$ , which *B*, given her posterior beliefs, will accept upon observing either ( $\{w, nk\} \vee \{s, nk\}$ ) or ( $\{s, k\}$ ). The risk that *A* runs, of course, is that *B* actually observes ( $\{w, k\}$ ) and hence knows that *A* is bluffing, in which case war occurs. But if the temptation is sufficiently large and the probability that *B* discovers the truth (that is,  $p_1$ ) sufficiently low, then *A* will be willing to take that risk and bluff. In that case, *A* deviates from the equilibrium path, and this can therefore not form the basis of a perfect Bayesian equilibrium. There is hence no separating equilibrium (details of the derivation and proofs in the appendix).

**Lemma 1.** Assume  $p^s - p^w > 2c / q_2^A$ , where  $q_2^A = p_2 / (p_1 + p_2)$ . Then there is no peaceful separating perfect Bayesian equilibrium.<sup>10</sup>

This is intuitive.  $q_2^A$  is *A*'s posterior probability of being at  $\{w, nk\}$  upon observing ( $\{w, k\} \vee \{w, nk\}$ ). A large  $q_2^A$  means a low probability of being caught deviating, and hence that *A* will be more willing to deviate.

### Pooling equilibrium

Consider now a pooling equilibrium in which both types of *A* offer  $x^*$ . Clearly,  $x^*$  must be large enough to satisfy a strong *A*, as otherwise *A* would prefer fighting to that agreement. But if  $x^*$  is too large, then *B* will reject it upon observing ( $\{w, nk\} \vee \{s, nk\}$ ), leading to war—an outcome that neither type of *A* would like.

**Lemma 2.** Assume  $\frac{2c}{q_2^B} > p^s - p^w > \frac{2c}{q_2^A}$ , where  $q_2^A = \frac{p_2}{p_1 + p_2}$  and  $q_2^B = \frac{p_2}{p_2 + p_3}$ . Then there exists a pooling perfect Bayesian equilibrium<sup>11</sup> in which all types of *A* offer  $x^* = p^s - c$ , which *B* accepts unless she observes ( $\{w, k\}$ ), in which case she rejects and war ensues.

The main proposition then follows immediately: for certain combinations of  $p^\phi$  and probability distributions on Nature's moves, there are no peaceful equilibria, and war occurs with positive probability.

<sup>10</sup>A Perfect Bayesian Equilibrium is a pair ( $s; b$ ) of strategy profile  $s$  and beliefs  $b$  such that  $s$  is sequentially rational given beliefs  $b$ , and  $b$  is consistent with  $s$ . See Fudenberg and Tirole (1991, 215) for a more formal definition.

<sup>11</sup>Note that this implicitly assumes that  $p_1 \leq p_3$ .

**Proposition 1.** Assume  $\frac{2c}{q_2^s} > p^s - p^w > \frac{2c}{q_2^k}$ . Then in every pure strategy Perfect Bayesian Nash equilibrium, war occurs with probability  $p_1$ .

This implies that even if both  $A$  and  $B$  know each other's attributes (that is,  $w, k$  is the true state of the world), war occurs in equilibrium.<sup>12</sup>

Could this inefficiency result be an artefact of the simple bargaining protocol? A first limitation of the one-shot bargaining model above is that it prevents learning. Repeating the game may, for example, allow  $B$  to reject without fearing immediate war, and as a result  $A$  might learn  $B$ 's type by varying her offers over time. This intuition turns out to be correct only under restrictive conditions. To see why, suppose first that the model is modified to include two periods, keeping only the two types of  $B$  as above—a knowledgeable and a non-knowledgeable one. In that case,  $A^w$  can make an offer in the first period such that only the non-knowledgeable  $B$  types would accept, and the knowledgeable types would accept in the second period. However this result is dependent on having only two types of  $B$ —a very restrictive assumption, as there is typically a continuum of types, and hence an infinite number of them. With two periods and three types, for example (such as knowledgeable, non-knowledgeable and partially knowledgeable types),  $A$  may not be able to separate all types. Instead, she may screen out one type in the first period, and be left with two types in the second period—a scenario equivalent to the one in the simple model above, so that war occurs with positive probability. This result is discussed formally in the online appendix.

The reader might think that, with an infinity of periods,  $A$  could progressively screen all types of  $B$ , until only the knowledgeable type would be left, to whom she could make a generous offer with probability one. This is not the case, however, because  $B^k$  (the knowledgeable type of  $B$ ) will not wait forever to be rewarded with that generous offer. If there are too many types to screen,  $B^k$  would rather fight today than wait for  $A$  to iteratively screen out all the other types. This result is discussed more formally in the online appendix.

As another improvement to the model, we could imagine that if  $B^k$  had the opportunity to tell  $A$  that he knows she is weak, then war might be avoided.  $B^k$  could, for example, make statements about  $A$ 's lack of resolve, poor troop training or high costs of war. However, introducing communication between the players has no effect on the equilibrium. Suppose, for example, that  $B$  could send a message immediately before  $A$ 's move. As long as the announcement has no direct effect on either player's payoff, then it is easy to show that  $B$ 's message also has no effect on the probability of war (see the online appendix for a more formal treatment). The logic is simple. A message could only lower the probability of war by increasing the probability that  $A$ , upon receiving that message, would make a generous offer to  $B^k$ . If so, however, all types of  $B$ —both knowledgeable and non-knowledgeable—would want to send that message. This means, however, that the message could not discriminate between types, and hence that it would not lead  $A$  to make a generous offer to its sender, contradicting the assumption above.

### The 2003 Iraq War

The negotiations over inspections in the lead-up to the 2003 Iraq War illustrate many aspects of this model. To simplify, Iraq ( $A$  in the model above) had two possible types: compliant (non-compliant) without (with) weapons of mass destruction (WMDs). There were also two types of opponent ( $B$ ): a knowledgeable (ignorant) one, who knew about (ignored) the absence of WMDs.

One interpretation is that  $B$  is Iran. Indeed, archives reveal that an important cause for Hussein's choices was his belief that Iran may have thought Iraq had WMDs. In terms of the model above, this is equivalent to  $A$  believing that  $B$  thinks she is strong and making high offers

<sup>12</sup>There also exist an infinity of semi-separating equilibria, each of which also involves a positive probability of war.

as a result (that is, refusing inspections).<sup>13</sup> Hussein himself explicitly stated this reason in his post-capture interrogations:

The threat from Iran was the major factor as to why he did not allow the return of the UN inspectors. Hussein stated he was more concerned about Iran's discovering Iraq's weaknesses and vulnerabilities than the repercussions of the United States for his refusal to allow UN inspectors back into Iraq.<sup>14</sup>

Another reading of the events is that *B* represents the United States. Saddam Hussein knew, for practical purposes, of American military superiority.<sup>15</sup> Moreover, there is substantial evidence to suggest that the United States also knew of Iraq's capabilities and (absence of) WMDs. Then CIA Director Tenet testified that 'We said that Saddam did not have a nuclear weapon and probably would have been unable to make one until 2007 to 2009'. In addition, a 1998 International Atomic Energy Agency (IAEA) report stated that, 'based on all credible information to date, the IAEA has found no indication of Iraq having achieved its program goal of producing nuclear weapons or of Iraq having retained a physical capability for the production of weaponisable nuclear material or having clandestinely obtained such material' (Pfiffner 2004).<sup>16</sup>

However, Hussein did not know how much exactly the United States knew—that is, whether he was in state  $\{w, k\}$  or  $\{w, nk\}$ . He had initially assumed that the United States must know that the evidence presented about the WMDs was erroneous (Duelfer and Dyson 2011). Yet statements by the US administration added to the confusion—for example when Bush declared in December 2002 that 'we do not know whether or not [Iraq] has a nuclear weapon'. This uncertainty about what the United States really believed led Hussein to refuse inspections and play tough, in the same way that *A* in the model above makes a large offer to *B*.<sup>17</sup>

Admittedly, the lead-up to the Iraq War of 2003 was a far more complex negotiation than the simple take-it-or-leave-it protocol can represent. It involved multiple players, and for Hussein a difficult balancing act of convincing the United States that he was not a danger, while making sure that Iran would think he was. Moreover, the negotiations were not just about a territory, but also about capabilities and knowledge, since what the US coalition was asking was to curtail Iraq's power, and to make her existing capabilities common knowledge by allowing inspections.<sup>18</sup> Nevertheless, the dynamic between the three countries illustrates some of the incentives and

<sup>13</sup>For a discussion of the role of Iran in Iraq's decision, see Lake (2010); for dissenting views, see Koblenz (2018). Note also that a similar argument can be applied to internal dissidents, who needed to be kept in fear of the existence of WMDs.

<sup>14</sup>US Federal Bureau of Investigation, *Casual Conversation with Saddam Hussein*, 11 June 2004. National Security Archive Electronic Briefing Book No. 279.

<sup>15</sup>Lake (2010, 31) suggests that the war may have occurred at least in part because of Iraq's uncertainty about US resolve. This does not contradict the argument here. Iraq's perception of US resolve depended on what the United States really knew or believed. Because Saddam did not really believe the United States would think the WMDs remained, he assumed the United States was posturing and had low resolve.

<sup>16</sup>This is of course a highly debated subject, well beyond the scope of this article. However, there is ample evidence that the intelligence community's information was, at the very least, far more nuanced than the White House administration's statements. Dick Cheney himself admitted that 'we never had any evidence that he had acquired a nuclear weapon', and Georg Thielmann, former director of the State Department Bureau of Intelligence and Research's program on strategic proliferation and military affairs, said: 'We never assessed that there was good evidence that Iraq was reconstituting or getting really serious about its nuclear weapons program' (Pfiffner 2004). For opposite views, see Debs and Monteiro (2014).

<sup>17</sup>Later, when it became clear that refusing inspections was a real *casus belli*, many of Saddam's actions then aimed at ensuring that the United States would not get the wrong impression, but he could never be sure what the Americans really believed (Woods et al. 2006, 93–4). He 'allowed the UN inspectors back into Iraq to counter allegations by the British Government. Hussein stated this was a very difficult decision to make, but the British Government had prepared a report containing inaccurate intelligence. It was this inaccurate intelligence on which the United States was making their decisions'. US Federal Bureau of Investigation, *Casual Conversation with Saddam Hussein*, 11 June 2004. National Security Archive Electronic Briefing Book No. 279.

<sup>18</sup>See Chadeaux (2011) for a discussion of negotiations about capabilities themselves.

strategies uncovered in the simple, ideal-type game presented above. Other examples below further illustrate the model.

### Reaching Common Knowledge

The model above showed that, to avoid war, countries must bridge their perceptions not only of each other's capabilities and resolve, but also of what each knows. Narrowing this gap, however, may prove more difficult than in the case of first-order asymmetric information. First, the need for higher-order information means that states not only need to send a message, but also to ensure that the message is received and correctly understood. A message—a 'confirmation'—is therefore also needed, which is itself subject to further miscommunication and misperception. A complication, however, is that states often have an incentive to misrepresent what they know. Finally, I discuss ways in which states may overcome some of these difficulties and create mechanisms and institutions to create common knowledge.

### Learning what the other knows

States acquire information about each other in two main ways: intelligence and signalling. Both can fail, however, and the mere possibility of these failures leads countries to always wonder: was my signal misunderstood? Did the other respond in that way because he does not know or because he knows and still chooses to make that offer? And what has he discovered about me?

#### *Intelligence*

Intelligence involves the use of sensors, spies or the decryption of messages to collect bits of information that may then be assembled into a coherent whole.<sup>19</sup> Intelligence often fails, however. Human intelligence relies on interrogations and conversations with those who have information, but may fail due to manipulation, false information being fed to the agent, or a simple inability to penetrate the centre of power (Ferris 2007). Other sources of failure include cognitive biases or organizational problems (Garthoff 2015; Shore 2005). Signals and imagery intelligence, which consist of information collected using radars, sonars or satellites, is limited to what is visible and tangible. Foliage alone may be sufficient to thwart the efforts of a drone or satellite. Finally, open source intelligence relies on the analysis of journals, radio and television, but is doubly limited: first by the information available to the other country's media, and second by possible manipulation.

Because these limitations of intelligence gathering also apply to the opponent, knowing how much the adversary knows is difficult. Has the other discovered my secret nuclear programme, or the absence thereof? Are spies providing information to the opponent? What is made of that information? And does the other believe it? Intelligence may also be manipulated to feed false information to the enemy. But did they believe that information? And did they think that I thought they believed it? The mere possibility of intelligence failures, in short, is sufficient to create uncertainty about what the other knows.

Diplomats spend considerable efforts analysing what their counterparts think, their perception of their own understanding, and so on. A comprehensive review of pre-World War I British and French diplomatic cables thus reveals that their concern was often to establish facts not about capabilities or resolve themselves, but rather about the other's thoughts, beliefs and estimates.<sup>20</sup>

<sup>19</sup>Most historical evidence on intelligence is limited to first-order intelligence – the collection of information about other actors' forces, strategy or motivation – but ignores higher-order intelligence. Yarhi-Milo (2013), for example, studies how leaders assess each other's intentions; May (2014) examines the shortcomings of intelligence before the world wars; Kahn (2000) and Ferris (2007) focus on intelligence during war.

<sup>20</sup>Pre-WWI cables have the distinct advantage of having been released in a comprehensive manner by most parties involved. E.g., Great Britain, Foreign Office, *British Documents On the Origins of the War (BDOOW)*, 1898–1914, London, 1926. France, *Ministère des affaires étrangères, Documents diplomatiques français. 1871–1914*.



For example a 1913 cable from Berlin reports, 'It is thought, or pretended to be thought, in some German circles connected with the Navy, that England's political affairs would render it difficult for any British Government to justify Naval increases'.<sup>21</sup>

The July 1914 crisis illustrates in a dramatic way the importance for diplomats of knowing what the other knows. Immediately after the assassination of Franz-Ferdinand, Austria asked Germany for her support in the Serbian crisis. In particular, Germany was informed of the demands that Austria would make from Serbia, and that these demands were purposefully harsh to elicit Serbia's refusal.<sup>22</sup> Wilhelm II, the German emperor, responded that 'Germany would support the Monarchy through thick and thin'.

Knowledge of Germany's knowledge of Austria's plan was crucial. If other capitals had known that Germany knew it, and yet had not pressured Austria to exercise restraint, then her 'revisionist' intentions would have been clearer. Thus, in a confidential cable to Foreign Secretary Edward Grey, the British ambassador notes: 'France and Britain are asking Germany to moderate Austria, but her intentions are unclear. They depend for example on what Germany knew'.<sup>23</sup> This knowledge of Germany's information was critical, as it would have facilitated Britain's choice to enter the war, and possibly would have deterred Germany in the first place.

Furthermore, Germany knew of Britain's ignorance and worked hard to make sure that British leaders would remain uninformed of Germany's knowledge: '[German] Secretary of State again denied that he had had any previous knowledge of terms of Austro-Hungarian note';<sup>24</sup> 'German Ambassador read me a telegram from German Foreign Office saying that Germany had not known beforehand and had had no more than other Powers to do with the stiff terms of Austrian note to Serbia' and 'Secretary of State again repeated very earnestly that he had had no previous knowledge of contents of Austro-Hungarian note, although he had been accused of knowing all about it'.<sup>25</sup> The French also did not know: 'M. Jules Cambon asked M. v. Jagow [*note: the German Secretary of State*] what were the terms of the Austrian note. The latter replied that he did not know'.<sup>26</sup> At the same time, European diplomats tried hard to understand how much Germany really knew: 'I am privately informed that German Ambassador knew text of Austrian note to Serbia before it was sent off and telegraphed it to the German Emperor, but I am not able to verify this, though I know from German Ambassador himself that he endorses every line of it'.<sup>27</sup>

A similar issue arose between Austria and Russia. On 17 July, Austria-Hungarian ambassador Szapáry 'expressed a desire to see Sazonov [Russian Foreign Minister] as soon as possible' (Schilling 1925, 26). As McMeekin (2013, 133–4) puts it, 'Berchtold [Foreign Minister of Austria-Hungary] wished to find out whether the Russians knew what he was up to'. In other words, Austria wanted to know how much Russia knew (second-order uncertainty). Furthermore, the Russians also wanted to ensure that the Austrians did not know that they knew (third-order uncertainty): 'Just as Berchtold wanted to be sure the Russians did not know what he was up to, so did the latter not want the Austrians to know that they were cottoning to the game' (McMeekin 1013, 134). As a result, Austria believed that it could finally draft the 48-hour

<sup>21</sup>Captain H. Watson to Sir E. Goschen. BDOOW, 20 January 1913.

<sup>22</sup>Evidence of this knowledge is evident, for example, in the following telegram of 14 July, from Vienna to Berlin: 'The ultimatum to be sent to Serbia "is being composed so that the possibility of its acceptance is practically excluded". But it has not yet been put in final form and would not be before July 19'. ('Too bad', notes the German Emperor). Quoted in Fromkin (2009, 166).

<sup>23</sup>Sir H. Rumbold to Sir Edward Grey. Berlin, 24 July 1914.

<sup>24</sup>Sir H. Rumbold to Sir Edward Grey. Berlin, 24 July 1914. On Germany's knowledge of Britain's ignorance, the German ambassador in London reported that he had discussed the Sarajevo aftermath and Austria's possible response with Sir Edward Grey. He stated that Grey 'was in a thoroughly confident mood, and declared in cheerful tones that he saw no reason for taking a pessimistic view of the situation' (9 July 1914).

<sup>25</sup>Sir H. Rumbold to Sir Edward Grey, 25 July 1914. BDOOW, Vol. XI, No. 157.

<sup>26</sup>Sir Edward Grey, 23 July 1914. BDOOW, Vol. XI, No. 87.

<sup>27</sup>Sir Edward Grey to Sir H. Rumbold, 25 July 1914. BDOOW, Vol. XI, No. 197.

ultimatum to Serbia since, 'as far as Berchtold knew [...], Austria had kept the other powers entirely in the dark [...]'. In fact, Berchtold continued to take on elaborate cover stories to ensure the secret of his planned ultimatum, unaware that the secret had been revealed to most European leaders by then.

### *Signalling*

States can also volunteer information by sending various forms of signals. A large literature in international relations has focused on this strategic exchange of information (Fearon 1997; Kydd 2005; Schelling 1980; Trager 2010), including the use of audience costs (Fearon 1994, but see also Snyder and Borghard 2011), or even cheap talk (Crawford and Sobel 1982; Farrell and Gibbons 1989; Sartori 2002; Sartori 2013; Trager 2010). These messages aim to convey information about capabilities and resolve, but also beliefs about those of the other. Yet, signals also often fail. They may fail to be detected or may be interpreted as informative signals (Jervis 2002; Mercer 2010) and may be affected by a variety of biases (Holsti 1962; Jervis 2015; Yarhi-Milo 2013). These well-known challenges associated with conveying information are amplified when dealing with common knowledge, because of the need for higher-order levels of information. A signal need not only be sent; confirmation of its reception and interpretation must also be conveyed back to the sender. The recipient might even have confirmed receipt, but not know whether the sender has received that receipt. This problem, known as the coordinated attack problem, illustrates one of the difficulties associated with reaching common knowledge.<sup>28</sup> There is also evidence of our cognitive limitations in dealing with higher-order beliefs (Nagel 1995). In particular, higher levels of the thought process are associated with more noise and suboptimal choices (Kübler and Weizsäcker 2004).

To summarize, the transmission of information often fails because of signal misinterpretation or intelligence failures. The very fallibility of these systems, or even the possibility that they might fail, means that states can hardly ever be certain of what their opponent knows, and common knowledge of information partitions is therefore unlikely. In addition, because the perception of what states know matters, states might not be willing to convey their private information about what they know, even if they could. Indeed, states are likely to be strategic with respect to the information they release about what they know of the adversary. Just as with first-order incomplete information, states have an incentive to misrepresent their higher-order private information.

At the beginning of the Cuban missile crisis, for example, President Kennedy did not want the Soviet Union to know what he knew about the missiles. When he met with Soviet Foreign Minister Andrei Gromyko on 18th October, he was assured that all Soviet assistance to Havana was only 'pursued solely for the purpose of contributing to the defense capabilities of Cuba'.<sup>29</sup> Kennedy did not then inform him of his knowledge. He continued to avoid any disruption to his public schedule to avoid suspicions that he might know. Later, once the Navy began to enforce the quarantine, however, communications began to be purposefully sent uncoded. The point then was to ensure that the Soviets would not misperceive their intentions, as well as to ensure their knowledge of the Americans' knowledge.

### *Creating common knowledge*

There are, however, ways in which states can convey information about attributes such as their military capabilities so that they become common knowledge. They may, for instance, choose to

<sup>28</sup>Suppose, for example, that *A* sends a credible message to *B*, but that with a small but positive probability the message will be garbled or misinterpreted. Even supposing that the message arrives properly, *B* still needs to make it clear to *A* that she received the message. But again, her confirmation and the interpretation might be garbled. So by now *A* knows, *B* knows, and *A* knows that *B* knows, but *B* does not know that *A* knows that *B* knows. So *A* needs to send a confirmation, and so on. See Rubinstein (1989) for a discussion of this problem.

<sup>29</sup>John F. Kennedy, 'Cuban Missile Crisis Address to the Nation', 22 October 1962.

reveal them publicly. Displays of artillery on national commemoration days, for example, or heavily publicized tests of new weapons thus not only signal strength, but also ensure that these capabilities are common knowledge. Joint military manoeuvres may also serve the same purpose: by granting the other a more intimate access to her resources, military capabilities become common knowledge.<sup>30</sup> Similarly, states may want to publicly convey their knowledge of the other's capabilities. They may, for example, mention the precise location of a nuclear site to convey their information in such a way that it becomes common knowledge. In the Cuban missile crisis, for example, satellite photos were used to show the United States' knowledge of the hidden ballistic missiles on Cuban soil.

Far more difficult, however, is to generate common knowledge of the *absence* of certain capabilities. Conveying positive knowledge—for example using public displays of force—is far easier than conveying negative knowledge—showing that certain capabilities do not exist. Saddam Hussein thus struggled to prove the absence of a WMD programme, whereas the Bush administration could easily pretend it believed Iraq had one. The willingness to co-operate with inspectors was a key signal in that regard, but came at a cost as the inspectors had to be trusted to be impartial and not to divulge other strategic information. Similarly, while public displays of force can easily create common knowledge that I have *at least*  $x$  units, showing that I have *no more* than  $z$  units is far more difficult. Public announcements are only helpful to reveal positive things that only a knowledgeable party would know.

## Conclusion

The role of information in bargaining outcomes has been a central theme in existing research on interstate conflict. However, models that incorporate asymmetric information have typically limited their attention to private information about fundamentals—first-order incomplete information—while assuming common knowledge of information partitions. Yet, assuming that states know whether their opponent knows the distribution of power or costs is implausible. The very uncertainty that is inherent to information collection—whether it be through intelligence or signalling—implies that states can rarely be certain of what their adversary managed to infer about them. States, in other words, are almost always uncertain of each other's uncertainty.

The importance of relaxing the assumption of common knowledge about information itself was illustrated using a simple bargaining model in which all equilibria included a positive probability of war. Because *A* thought that *B* might think she was strong, *A* had an incentive to bluff and demand large concessions from *B*. With positive probability, however, *B* was well aware of *A*'s weakness, and hence rejected the offer in favour of war. The results have important implications for our understanding of the conditions under which war can emerge, for the way in which we should model uncertainty, and for our empirical approach to testing the role of uncertainty in international negotiations.

The importance of higher-order uncertainty also implies that states might behave strategically with regards to how they release or hide information about what they know of the enemy (Prunckun 2012, Ch. 8). Classified information, for example, often hides not only what you have, but also what you know about the other. Similarly, credible signalling is not limited to conveying information about your capabilities and resolve, but also includes informing (or not) the other about what you know of him. Just as in the case of first-order intelligence, states have an incentive to misrepresent their beliefs about the other and to deceive them about what they know and what they do not know. Depending on their needs, they might exaggerate their perception of the other's capabilities, or downplay them.

The results also have important empirical implications. Inferences about the role of private information may be incorrect if they ignore higher-order information—whether they rely on

<sup>30</sup>For more on the creation of common knowledge, see also Morrow (2014).

historical case studies or large-N analyses. Two otherwise identical cases with the same first-order information structure may result in opposite outcomes depending on higher-order beliefs. For example, before launching the surprise attack that would start the Yom Kippur War, Egypt's Sadat sent scouts across the border to ensure that the Israelis had not discovered the pre-emptive strike they were planning—that is, to ensure that they indeed did not know what he knew. But consider the hypothetical case in which Sadat had erroneously thought that Israel knew about the attack, and hence cancelled his surprise attack. In that case the inference using data on first-order information only (that is, that Israel did not know) would have been that incomplete information does not correlate with war onset. Only by incorporating second-order information would the correct inference have been drawn.<sup>31</sup>

Explicitly modelling higher-order uncertainty leads to a more refined understanding of strategic interactions, with potentially important consequences for the efficiency of bargaining. If states need to know what each other knows, then the conditions for peace are more stringent than the simple exchange of information about first-order attributes such as capabilities or resolve. More generally, modelling larger information structures may lead to a deeper understanding of states' strategies and more a subtle understanding of their communication. Future research will need to focus on the mechanisms states use to address the problems associated with higher-order uncertainty. International organizations, mediation and monitoring, in particular, may all play a significant role in the creation of common knowledge.

**Acknowledgements.** I thank Constantine Boussalis, Gail McElroy, James Morrow, William Phelan, Anne Sartori, Peter Stone, Richard Van Weelden, George Yin and anonymous reviewers for helpful comments. Fiona May and Christian Oswald provided exceptional research assistance.

## References

- Acharya A and Grillo E (2015) War with crazy types. *Political Science Research and Methods* 3 (2):281–307.
- Aumann RJ (1976) Agreeing to disagree. *The Annals of Statistics* 4 (6):1236–1239.
- Bas M and Schub R (2016a) How uncertainty about war outcomes affects war onset. *Journal of Conflict Resolution* 60 (6):1099–1128.
- Bas M and Schub R (2016b) Mutual optimism as a cause of conflict: secret alliances and conflict onset. *International Studies Quarterly* 60 (3):552–564.
- Blainey G (1988) *Causes of war*. New York: Simon and Schuster.
- Chadeaux T (2011) Bargaining over power: when do shifts in power lead to war? *International Theory* 3 (2):228–253.
- Chadeaux T (2012) War as an investment. Unpublished manuscript.
- Crawford VP and Sobel J (1982) Strategic information transmission. *Econometrica* 50 (6):1431–1451.
- Debs A and Monteiro NP (2014) Known unknowns: power shifts, uncertainty, and war. *International Organization* 68 (1):1–31.
- Duelfer CA and Dyson SB (2011) Chronic misperception and international conflict: the US-Iraq experience. *International Security* 36 (1):73–100.
- Farrell J and Gibbons R (1989) Cheap talk can matter in bargaining. *Journal of Economic Theory* 48 (1):221–237.
- Fearon JD (1994) Domestic political audiences and the escalation of international disputes. *American Political Science Review* 88 (3):577–592.
- Fearon JD (1995) Rationalist explanations for war. *International Organization* 49 (3):379–414.
- Fearon JD (1997) Signaling foreign policy interests: tying hands versus sinking costs. *Journal of Conflict Resolution* 41 (1):68–90.
- Ferris J (2007) *Intelligence and Strategy: Selected Essays*. Abingdon: Routledge.
- Fey M and Ramsay KW (2007) Mutual optimism and war. *American Journal of Political Science* 51 (4):738–754.
- Filson D and Werner S (2002) A bargaining model of war and peace: anticipating the onset, duration, and outcome of war. *American Journal of Political Science* 46 (4):819–837.

<sup>31</sup>Note that the uncertainty in this case is about Sadat's strategy, and not a parameter of the game, in contrast to the model above. However, surprise attacks are akin to additional power in the game, and hence not all that different from uncertainty about capabilities. I thank an anonymous reviewer for this comment.

- Fromkin D** (2009) *Europe's Last Summer*. New York: Random House.
- Fudenberg D and Tirole J** (1991) *Game Theory*. Cambridge, MA: MIT Press.
- Garthoff RL** (2015) *Soviet Leaders and Intelligence: Assessing the American Adversary During the Cold War*. Washington, DC: Georgetown University Press.
- Gartzke E** (1999) War is in the error term. *International Organization* 53 (3):567–587.
- Harsanyi JC** (2004) Games with incomplete information played by Bayesian players, I–III: Part I. The basic model. *Management Science* 50 (12):1804–1817.
- Holsti OR** (1962) The belief system and national images: a case study. *Journal of Conflict Resolution* 6 (3):244–252.
- Jervis R** (1976) *Perception and Misperception in International Politics*. Princeton, NJ: Princeton University Press.
- Jervis R** (2002) Signaling and perception: drawing inferences and projecting images. In Monroe K, (ed.), *Political Psychology*. Hillsdale, NJ: Lawrence Erlbaum, pp. 293–314.
- Jervis R** (2015) *Perception and Misperception in International Politics*. Princeton, NJ: Princeton University Press.
- Kahn D** (2000) *Hitler's Spies: German Military Intelligence in World War II*. Cambridge, MA: Da Capo Press.
- Koblentz GD** (2018) Saddam versus the inspectors: the impact of regime security on the verification of Iraq's WMD disarmament. *Journal of Strategic Studies* 41 (3):372–409.
- Kübler D and Weizsäcker G** (2004) Limited depth of reasoning and failure of cascade formation in the laboratory. *The Review of Economic Studies* 71 (2):425–441.
- Kydd AH** (2005) *Trust and Mistrust in International Relations*. Princeton, NJ: Princeton University Press.
- Lai B** (2004) The effects of different types of military mobilization on the outcome of international crises. *Journal of Conflict Resolution* 48 (2):211–229.
- Lake DA** (2010) Two cheers for bargaining theory: assessing rationalist explanations of the Iraq war. *International Security* 35 (3):7–52.
- May ER** (2014) *Knowing one's enemies*. Princeton, NJ: Princeton University Press.
- McMeekin S** (2013) *July 1914 Morrow. Countdown to War*. New York: Basic Books.
- Mercer J** (2010) Emotional beliefs. *International Organization* 64 (1):1–31.
- Milgrom P and Stokey N** (1982) Information, trade and common knowledge. *Journal of Economic Theory* 26 (1):17–27.
- Morrow JD** (1989) Capabilities, uncertainty, and resolve: a limited information model of crisis bargaining. *American Journal of Political Science* 33 (4):941–972.
- Morrow JD** (2014) *Order Within Anarchy: The Laws of War as An International Institution*. Cambridge: Cambridge University Press.
- Nagel R** (1995) Unraveling in guessing games: an experimental study. *The American Economic Review* 85 (5):1313–1326.
- Pfiffner JP** (2004) Did President Bush mislead the country in his arguments for war with Iraq? *Presidential Studies Quarterly* 34 (1):25–46.
- Powell R** (1996) Bargaining in the shadow of power. *Games and Economic Behavior* 15 (2):255–289.
- Powell R** (2004a) Bargaining and learning while fighting. *American Journal of Political Science* 48 (2):344–361.
- Powell R** (2004b) The inefficient use of power: costly conflict with complete information. *American Political Science Review* 98 (2):231–241.
- Prunckun H** (2012) *Counterintelligence Theory and Practice*. Lanham, MD: Rowman & Littlefield.
- Reed W** (2003) Information, power, and war. *American Political Science Review* 97 (4):633–641.
- Reiter D** (2003) Exploring the bargaining model of war. *Perspective on Politics* 1 (1):27–43.
- Rubinstein A** (1989) The electronic mail game: strategic behavior under 'almost common knowledge'. *The American Economic Review* 79 (3):385–391.
- Sartori AE** (2002) The might of the pen: a reputational theory of communication in international disputes. *International Organization* 56 (1):121–149.
- Sartori AE** (2013) *Deterrence by Diplomacy*. Princeton, NJ: Princeton University Press.
- Schelling TC** (1980) *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.
- Schilling MF** (1925) *How the War Began in 1914*. London: George Allen.
- Schub R** (2015) Are you certain? Leaders, overprecision, and war. Unpublished manuscript.
- Schultz KA** (1998) Domestic opposition and signaling in international crises. *American Political Science Review* 92 (4):829–844.
- Shore Z** (2005) *What Hitler Knew: The Battle for Information in Nazi Foreign Policy*. Oxford: Oxford University Press.
- Slantchev BL** (2003) The principle of convergence in wartime negotiations. *American Political Science Review* 97 (4):621–632.
- Slantchev BL** (2004) How initiators end their wars: the duration of warfare and the terms of peace. *American Journal of Political Science* 48 (4):813–829.
- Slantchev BL and Tarar A** (2011) Mutual optimism as a rationalist explanation of war. *American Journal of Political Science* 55 (1):135–148.
- Smith A and Stam A** (2004) Bargaining and the nature of war. *Journal of Conflict Resolution* 48 (6):783–813.

**Snyder J and Borghard ED** (2011) The cost of empty threats: a penny, not a pound. *American Political Science Review* **105** (3):437–456.  
**Toft MD** (2006) Issue indivisibility and time horizons as rationalist explanations for war. *Security Studies* **15** (1):34–69.  
**Trager RF** (2010) Diplomatic calculus in anarchy: how communication matters. *American Political Science Review* **104** (2):347–368.  
**Wagner R** (2000) Bargaining and war. *American Journal of Political Science* **44** (3):469–484.  
**Wittman D** (1979) How a war ends: a rational model approach. *Journal of Conflict Resolution* **23** (4):743–763.  
**Woods KM, Pease MR, Stout ME, Murray W and Lacey JG** (2006) Iraqi perspectives project. A view of operation Iraqi freedom from Saddam’s senior leadership. Technical report, United States Joint Forces Command.  
**Yarhi-Milo K** (2013) In the eye of the beholder: how leaders and intelligence communities assess the intentions of adversaries. *International Security* **38** (1):7–51.

## Appendix

### Proofs

*Proof of Lemma 1 (no separating equilibrium).* Consider first the possibility of a separating equilibrium in which *A* offers  $x^L$  upon observing  $(\{w, k\} \vee \{w, nk\})$  but  $x^H$  upon observing  $(\{s, nk\} \vee \{s, k\})$ . *B*’s posterior probabilities  $q_i$  of being at state  $s_i$  upon observing offer  $x^k$  and information set  $(\theta)$  are then:

$$\begin{cases} q_1^B = q(w, k \mid x^L, (\{w, k\})) = 1 \\ q_2^B = q(w, nk \mid x^L, (\{w, nk\} \vee \{s, nk\})) = 1 \\ q_3^B = q(s, nk \mid x^H, (\{w, nk\} \vee \{s, nk\})) = 1 \\ q_4^B = q(s, k \mid x^H, (\{s, k\})) = 1 \end{cases}$$

First note that if there is to be a peaceful equilibrium, then it must be that

$$\begin{cases} p^w - c \leq x^L \leq p^w + c \quad \text{and} \\ p^s - c \leq x^H \leq p^s + c \end{cases}$$

Let  $x^L = p^w - c + \alpha$  and  $x^H = p^s - c + \beta$ , where  $0 \leq \alpha, \beta \leq 2c$ . Clearly, given her beliefs, *B* always accepts *A*’s offer. However, note that *A*’s utility from playing  $(x^L, x^H)$  is  $p^w - c + \alpha$ . So *A*’s will deviate to an offer  $x^H$  instead of  $x^L$  upon observing  $(\{w, k\} \vee \{w, nk\})$  if

$$\begin{aligned} q_1(p^w - c) + (1 - q_1)(p^s - c + \beta) &> p^w - c + \alpha \\ \Rightarrow p^s - p^w &> \frac{\alpha(p_1 + p_2)}{p_2} - \beta. \end{aligned}$$

But we assumed that  $p^s - p^w > \frac{2c(p_1 + p_2)}{p_2}$ , which is clearly greater than or equal to  $\frac{\alpha(p_1 + p_2)}{p_2} - \beta$  for any  $0 \leq \alpha, \beta \leq 2c$ . This means that *A* will deviate (regardless of *B*’s beliefs off the equilibrium path), and hence that there can be no separating equilibrium.

*Proof of Lemma 2 (pooling equilibrium).* Clearly there can be no pooling equilibrium in which both types of *A* offer  $x^* < p^s - c$ , as a strong *A* would prefer war, and hence deviate. So consider a pooling equilibrium in which *A* offers  $x^* = p^s - c$ . In a pooling equilibrium, *B*’s posteriors are dictated by Bayes’ rule:

$$\begin{cases} q_1^B = q(w, k \mid x^*, (\{w, k\})) = p_1 \\ q_2^B = q(w, nk \mid x^*, (\{w, nk\} \vee \{s, nk\})) = \frac{p_2}{p_2 + p_3} \\ q_3^B = q(s, nk \mid x^*, (\{w, nk\} \vee \{s, nk\})) = \frac{p_3}{p_2 + p_3} \\ q_4^B = q(s, k \mid x^*, (\{s, k\})) = p_4 \end{cases}$$

*B*’s beliefs off the equilibrium path must be such that she assumes from any offer  $x' \neq x^*$  that she is facing a weak *A*. If not, then at least *A*’s will want to deviate from offering  $x^*$ , and there can hence be no equilibrium in which both types of *A* offer  $x^*$ .

Consider now what *B* will do upon observing  $x^*$ . At  $(\{w, k\})$ , *B* will reject because  $x^* = p^s - c > p^w + c$ . At  $(\{s, k\})$ , *B* always accepts. Finally at  $(\{w, nk\} \vee \{s, nk\})$ , *B* accepts  $x^* = p^s - c$  if the utility of accepting is greater than the utility of fighting weighted by the probability of the different types of *A*, i.e., if

$$\begin{aligned} 1 - x^* &\geq q_2^B(1 - p^w - c) + (1 - q_2^B)(1 - p^s - c) \\ \Rightarrow p^s - p^w &\leq \frac{2c(p_2 + p_3)}{p_2} \end{aligned}$$

which has been assumed.

Does  $A$  have an incentive to deviate from this strategy? Clearly, a strong  $A$  will not deviate, given  $B$ 's off the equilibrium path beliefs. But would a weak  $A$  deviate?  $A^w$  would deviate to offering  $p^w + c$  (any higher offer will be rejected, and a lower offer is clearly not rational) if:

$$\begin{aligned} p^w + c &> q_1^A(p^w - c) + (1 - q_1^A)(p^s - c) \\ \Rightarrow \frac{2c(p_1 + p_2)}{p_2} &> p^s - p^w \end{aligned}$$

But  $p^s - p^w > \frac{2c(p_1 + p_2)}{p_2}$  was assumed, so  $A$  would not deviate and hence this constitutes an equilibrium.

*Proof of Proposition 1.* The proof follows directly from Lemmas 1 and 2. To show that war occurs with positive probability in every PBE, simply replace  $x^* = p^s - c$  in the proof of lemma 2 with  $x^* = p^s - c + \gamma$ , where  $\gamma \in [0, 2c]$  is such that

$$\frac{2c - \gamma}{q_2^B} \geq p^s - p^c \geq \frac{2c}{q_2^A} - \gamma.$$