Are You Certain? Leaders, Overprecision, and War*

Robert Schub†

August 2015

Abstract

Does greater certainty about an opponent's military and political characteristics cause peace? This paper argues that the relationship between certainty and war is conditional on certainty's sources. Consistent with bargaining models of war, certainty leads to peace when it is appropriate given the available information. However, certainty often arises due to perceptual errors of leaders and advisors who overestimate the precision of available information. These errors of overprecision are common when decision-makers downplay the importance of political, as opposed to military, factors in determining conflict outcomes, which is likely when estimation processes marginalize elites who oversee a state's international diplomacy. In such cases, greater certainty can increase the likelihood of war. I test the hypotheses by measuring certainty in an original data set consisting of over 1,100 texts from declassified documents pertaining to 61 crises involving the US during the Cold War. Results support the theory, demonstrating the conditional relationship between certainty and war. The findings highlight the benefits of integrating behavioralist and rationalist approaches to improve our understanding of interstate conflict and flag policy conditions most conducive to perceptual errors that increase the risk of war.

*Department of Government, Harvard University. 1737 Cambridge St, Cambridge, MA 02138. Web: http://scholar.harvard.edu/schub. Email: rjschub@fas.harvard.edu.

†I am grateful to Erin Baggott, Muhammet Bas, Tyson Belanger, Julie Faller, Jeff Frieden, Jeff Friedman, Michael Gill, Nils Hagerdal, Michael Horowitz, Iain Johnston, Josh Kertzer, Noah Nathan, Rich Nielsen, Steve Rosen, Anne Sartori, Beth Simmons, Dustin Tingley, Ariel White, George Yin, Yuri Zhukov and participants at various Harvard workshops, the ISA and MPSA annual conferences, and the Empirical Implications of Bargaining Models Workshop at PSSI for helpful discussions and feedback.

1 Introduction

The latent "power to hurt" through military force shapes international politics by dictating which peaceful settlements are more attractive than war (Schelling, 1966, p. 2). However, precisely how much "hurt" states can impose or tolerate, and thus how unattractive war is, is frequently uncertain. This uncertainty has many sources. States often have limited information about the quantity and quality of an opponent's soldiers and arms. For instance, France's rationale for the Maginot Line hinged upon an incomplete understanding of Germany's inter-war developments. An opponent's resolve or willingness to bear costs is unclear, particularly in the absence of conflict itself. The US consistently underestimated North Vietnam's cost tolerance. Seven years after US combat troops arrived, Chairman of the Joint Chiefs Moorer was still surprised that the "North Vietnamese seem to be able to take unlimited losses."[1] Leaders struggle to understand an opponent's pre and post-conflict political landscape which affects the type of war and enemy to be fought. Before the Iraq War, the Bush administration expected a short conventional war against national forces, not the protracted counter-insurgency that emerged.

Facing such complexity, what determines how uncertain leaders are? How does the degree of uncertainty affect whether war occurs? To answer these questions, this paper develops and tests a theory of the causes and consequences of certainty in the decision for war. I argue that the relationship between certainty and war is conditional on certainty's sources. Certainty is pacifying when it is warranted given the information available. However, certainty can be unwarranted as leaders and advisors often believe their information is more precise than it actually is. I theorize that these errors of overprecision are most common when elites downplay the importance of political, as opposed to military, factors when estimating conflict outcomes. This is most likely when assessment processes marginalize a state's diplomats. When marginalized, greater certainty can increase the likelihood of war. To test the hypotheses I develop an original data set measuring elite

[1] Minutes of a Washington Special Actions Group Meeting, May 1, 1972, *FRUS*, Volume XIII, Vietnam, 1969-1976, Document 107.

certainty in declassified documents pertaining to 61 opportunities for the US to use force during the Cold War. Results strongly support the theoretical contention, demonstrating the conditional relationship between certainty and war.

Existing rationalist accounts suggest certainty fosters peace.[2] Reaching diplomatic settlements hinges upon the ability of actors with conflicting preferences to identify negotiated outcomes that are preferable to suffering the costs of conflict (Schelling, 1966; Blainey, 1988; Fearon, 1995; Powell, 1999; Fey and Ramsay, 2011; Slantchev and Tarar, 2011). Uncertainty arising from private information obscures an opponent's expected war outcome and thus the range of bargaining settlements that are preferable to fighting. If we assume actors are privy to accurate assessments of the degree of uncertainty present, greater certainty reduces the likelihood of war.

This framework assumes that actors are only certain when they have sufficient information to justify such certainty. In practice, however, estimative certainty is endogenously determined by the actors themselves through subjective processes. Theoretical and empirical behavioralist research points to a number of ways in which individual behavior systematically deviates from rationalist predictions, especially in the realm of information processing (for example, Simon (1955); Kahneman (2011)). A notable and consistent bias is perceptual error that causes individuals to overstate their degree of certainty (Alpert and Raiffa, 1982; Teigen and Jorgensen, 2005; Moore and Healy, 2008). Despite the elevated stakes associated with conflict, I demonstrate that elites are susceptible to making these errors of overprecision.

If certainty is unwarranted due to overprecision it can increase the ex ante probability of war.[3] I specify when certainty is likely to arise from perceptual error, rather than from quality information.

[2]This paper is concerned with informational mechanisms that cause war as opposed to commitment problems due to shifting power. Certainty is not inherently pacifying in the latter case.

[3]Unwarranted certainty due to overprecision implies a baseline 'truth' exists against which to compare estimates. Identifying overprecision ex ante is problematic as the true distribution is unknown. Even ex post this task might be impossible as revealed outcomes that differ from estimate means could be indicative of a bad draw rather than a flawed assessment. Though others propose strategies to work around this dilemma (Johnson and Tierney, 2011, footnote 7), I am skeptical about their feasibility. Instead, I posit causes of greater certainty and theorize the conditional effect of more certainty arising from specific causes on conflict propensity. As elaborated upon later, this design focuses on specifying the differential consequences of certainty rather than direct assessments of overconfidence.

Overprecision occurs when assessment processes marginalize advisors tasked with analyzing the political attributes of opponents—such as their resolve or post-conflict political landscapes—that affect conflict outcomes.

Bargaining theory's insight linking asymmetric information to war has achieved widespread and growing prominence in contemporary accounts of conflict. Yet the empirical implications have eluded systematic testing as measuring certainty poses a challenge to scholars (Gartzke and Poast, N.d.). I address the challenge by developing a measure of how US presidents and senior advisors evaluated the certainty attendant to possible conflicts during the Cold War. Focusing on elites captures estimated uncertainty at the locus of foreign policy decision-making. The measurement is based on over 1,100 declassified meeting transcripts and memoranda that were selected and pruned to match the theoretical concepts of interest. Consistent with the theory, when there is high State Department involvement certainty leads to peace. In the converse conditions, the relationship flips.

The findings highlight policy conditions in which leaders are prone to estimative errors that increase the probability of conflict. Additionally, they illustrate benefits to incorporating behavioralist insights into rationalist models. This is not a repudiation of the bargaining account of war. Rather, it is an effort to account for and specify when individual behavior is (un)likely to approximate rationalist predictions. Doing so is in the spirit of Fearon's (1995, p. 409) conclusion that after developing rationalist mechanisms for war "bounded rationality may appear a more important cause of disagreement" and follows a suggestion from Lake (2010) that "[a] marriage of behavioral and bargaining theory promises to be more powerful than either alone."

The remainder of the paper proceeds as follows. Section two discusses the existing literature on uncertainty's role in international conflict, its possible sources, and its effect on conflict likelihood. Section three develops the theory, beginning with the consequences of certainty and then turning to the causes which condition the consequences. Section four lays out the empirical framework and the operationalization of the uncertainty variable. Sections five and six empirically test the theory and assess robustness. Section seven concludes.

2 Incomplete Information and Certainty: Extant Accounts

Uncertainty is central to theories of international conflict.[4] Within the bargaining theory of war, information asymmetries generate uncertainty about opponent's capabilities or costs of war which can lead to the socially inefficient outcome of conflict (Fearon, 1995; Powell, 1999; Gartzke, 1999; Smith and Stam, 2006; Slantchev and Tarar, 2011; Fey and Ramsay, 2011). Alternatively, uncertainty can pertain to state intentions (Jervis, 1976; Kydd, 1997; Montgomery, 2006; Yarhi-Milo, 2013) or war's outcome if the fighting process is stochastic (Waltz, 1979; Mearsheimer, 2001; Bas and Schub, forthcoming). The focus here is on the first form pertaining to uncertainty over enemy capabilities and resolve.

Limited Information, Subjective Probabilities, and Certainty

Before addressing the international conflict literature, I clarify terms and review findings in studies of subjective probability. Consider a probability distribution for a variable of interest that affects bargaining prospects, such as costs or casualties in conflict. Individuals with limited information—such as leaders making decisions about the use of force—must formulate their own subjective distributions for this variable.[5] For instance, will the enemy capitulate after suffering 100 fatalities, 100,000 fatalities, or anything in between? This subjective assessment can be described by its variance and its mean. The former summarizes the uncertainty of the estimate and is the central concern for this paper. An estimate's uncertainty is a function of both the estimator's information

[4] I use the terms uncertainty and risk in a non-Knightian sense. That is, scenarios with probabilities less than one assigned to outcomes entail uncertainty and therefore risk. The probabilities may themselves be subjective—indeed, understanding biases in these subjective assessments is one goal of the paper—but are not intrinsically unknowable.

[5] It is unlikely decision-makers conceive of problems in terms of continuous distributions with exact means and variances. However, any estimate, whether of an opponent's military capabilities or of a war's outcome, implicitly constitutes a probability distribution. Consider assessments of how the Iraq War would unfold. Though some anticipated a protracted conflict (Shanker, 2007; Bennett and Stam, 2006), these views were largely downplayed. Secretary of Defense Rumsfeld expected the war "could last six days, six weeks. I doubt six months" (Rumsfeld, 2003). This constitutes a probability distribution that perhaps had a long tail—General Shinseki's estimate was widely cited—but where the vast majority of the likelihood is placed around the short duration side of the outcome spectrum, rendering a high certainty assessment. All estimates entail probability distributions, and therefore a measure of uncertainty.

level and an estimator's perceptual errors or biases.

Prior work in psychology and behavioral economics illustrates that errors of overprecision are widely prevalent when estimating the certainty of an assessment (Tversky, 1974; Alpert and Raiffa, 1982). When individuals are asked for an estimate with 90% confidence intervals, these intervals only contain the true answer 30% to 50% of the time (Alpert and Raiffa, 1982; Teigen and Jorgensen, 2005; Moore and Healy, 2008). Moreover, overprecision is not subject to reversals which can be elicited with other forms of overconfidence.[6] When errors of precision occur, they are excessively certain with single-point estimates representing the extreme case. Put simply, "(almost) everyone exhibits overprecision (almost) all the time" (Ortoleva and Snowberg, 2015).

Sources of Certainty

Certainty is a function of information and perceptual error. Rationalist accounts of war typically preclude the possibility of error and assume that actors know the probability distribution for variables of interest such as an enemy's resolve or its capabilities. While a state might invest in intelligence gathering and thus get more information about the distribution (Arena and Wolford, 2012), such a framework treats the resulting distribution as objective. Some of the conflict literature incorporates insights from cognitive psychology, such as an array of heuristics and shortcomings that affect decision making (Jervis, 1976; Biddle, 2004; Mercer, 2005; Rathbun, 2007; Kahneman and Renshon, 2007; Johnson and Tierney, 2011). However, few studies of conflict address what conditions elevate the likelihood of these biases emerging, instead treating these influences as constants

[6]Moore and Healy (2008) show overconfidence is an ill-defined term that conflates multiple phenomena including over-optimism and overprecision. The variance reflects an estimate's precision, a distribution's mean reflects an estimate's optimism. While optimism is critically important, I put aside its causes and consequences in favor of studying certainty, which has a far less developed literature pertaining to international security and has less immediate implications. This separation is justifiable because over-optimism and overprecision do not imply one another (Moore and Healy, 2008; Ortoleva and Snowberg, 2015). On optimism generally, see for example Tversky and Kahneman (1982) and Kahneman (2011). On optimism in conflict, see Johnson (2004); Kahneman and Renshon (2007); Johnson and Tierney (2011) among others.

that can explain why war might occur at all but not why it might occur in any specific instance.[7]

Results of Certainty

A more developed literature addresses uncertainty as a cause of conflict. Asymmetric information coupled with incentives to misrepresent is considered one of the primary causes of war. Information asymmetries can pertain to levels of resolve or willingness to bear costs (Fearon, 1994, 1995; Schultz, 1999), or capabilities affecting the probability of victory in conflict (Blainey, 1988; Reed, 2003; Smith and Stam, 2006; Slantchev and Tarar, 2011; Fey and Ramsay, 2011). Despite its theoretical prominence, empirically testing the relationship between asymmetric information and conflict remains problematic (Gartzke (1999), but see Bas and Schub (N.d.); Kaplow and Gartzke (2013)). In the spirit of Reed (2003), I develop empirical implications about the effect of uncertainty on the likelihood of conflict by allowing the magnitude of uncertainty to vary.

Less attention is devoted to subjective and potentially biased estimation in conflict models. Kurizaki (N.d.) allows for perceptual error in a crisis bargaining model in which states misperceive whether an opponent is issuing a threat. This approach leaves the sources of misperception unspecified and focuses on threat estimation rather than estimation pertinent to bargaining.[8] Mitzen and Schweller (2011) emphasize misplaced certainty due to affective sources as a cause of conflict that operates through security dilemmas as opposed to bargaining breakdowns. I similarly theorize the pernicious consequences of misplaced certainty, though the error emanates from a different source and conflict follows through a different logic.

[7] Beyond cognitive factors, prior work posits bureaucratic factors—such as a cult of the offensive, groupthink, and dysfunctional civil-military relations—shape perceptions of war outcomes (Allison, 1971; Janis, 1982; Kier, 1995; Van Evera, 1999; Feaver, 1999; Badie, 2010).

[8] Other studies address subjective estimation by endogenizing beliefs through inter-temporal preference shifts (Minozzi, 2013) or through non-common priors (Smith and Stam, 2004).

3 Theorizing Certainty's Conditional Effects

Consequences of Certainty

Does certainty cause war or peace? Consider a canonical model of war as a point of departure (Fearon, 1995). Two states A and B bargain over the allocation of a contested continuously divisible resource R, such as territory or a slate of policies, with a value normalized to equal one. For expositional clarity I adopt the simplest bargaining protocol. In a non-iterated interaction, A makes proposals (retaining x of 1) and B either accepts or rejects with the latter resulting in war. War is a costly winner-take-all lottery with A winning with probability p. Both sides suffer war costs (arising from casualties, destruction of the contested resource, or a negative reputation hit), c_A and c_B respectively, if conflict occurs.

With asymmetric information, B's reservation value—its expected payoff from war—is unknown to the proposing state. The proposing state's uncertainty about its enemy's expected war outcome can arise from uncertainty about the latter's resolve or war costs (c_B) or its military capabilities which affect the conflict outcome (p). Placing uncertainty over either parameter has similar implications for my purposes (though see Fey and Ramsay (2011)). I model the uncertainty over the opponent's resolve. When excluding the possibility of perceptual errors, uncertainty indicates that B's resolve falls somewhere in a known range of possible levels given by a prescribed cumulative distribution function $F(z)$ which is continuous.[9] More concretely, this could mean a state is unsure of an enemy's tolerance for fatalities. The well-known result from this framework is that the proposing state A balances a risk-return tradeoff, which leaves some positive ex ante probability that the enemy B rejects the offer and war occurs. For instance, the proposing state may make offers that satisfy an enemy with a moderate tolerance for battlefield casualties but leaves the risk of war if the enemy is willing to suffer the costs of total war. A bargain satisfying the moderately resolved enemy is preferable for the proposer compared to bargains satisfying a highly committed

[9] Results are no different with a discrete distribution but the intuition is less clear.

opponent but leaves a higher risk of war—hence the risk-return tradeoff.

When the distribution of possible enemy resolve levels is fixed and known, war occurs when the opponent happens to be a strong type, which is modeled as random and impedes empirical assessment of this central explanation for war. I follow Reed (2003) in deriving a result which provides some leverage on the problem. Rather than assuming enemy resolve is drawn from a fixed distribution $F(z)$, allow the variance of the distribution to vary while holding the mean constant.[10] That is, uncertainty differs across cases. Rather than having uncertainty over an opponent's resolve such that c_B is distributed, say, uniformly between 0 and 1—$c_B \sim U(0,1)$—imagine that greater intelligence allows the assessor to accurately restrict this such that $c_B \sim U(0.2,0.8)$. Using a beta distribution and reducing its variance produces substantively similar results.[11]

Varying the amount of certainty—that is, the variance—reveals that the probability of war is decreasing in certainty. Low certainty gives rise to the risk-reward tradeoff and attendant positive probability of war. Increased certainty minimizes the tradeoff and converges to the complete information case where the probability of war is zero. Figure 1's lower line plots the comparative static with c_B converging from $c_B \sim U(0.05,0.55)$ to $c_B = 0.3$ with $c_A = 0.2$ and $p = 0.4$.[12]

> *Hypothesis 1a: The ex ante probability of conflict declines as the certainty of the assessment increases, provided there is no perceptual error.*

Departing from the traditional rationalist framework, I now introduce perceptual error in the form of unwarranted certainty. Estimates are not objectively given but arise from subjective processes that are prone to unwarranted certainty (overprecision). Recognizing this subjectivity shifts the question of when war occurs from the realm of randomness—do we happen to face a strong

[10] Holding the mean constant best approximates different crises with varying degrees of certainty rather than a fixed crisis with learning occurring. As in Reed (2003), a fixed mean makes the model ill-suited for capturing learning dynamics except in the narrow case where signals of type all match the mean value of the initial distribution.

[11] As in Fearon (1995), the results are restricted to distributions with nondecreasing hazard rates $\frac{f(z)}{1-F(z)}$, which holds for all uniform distributions and all beta distributions with shape parameters greater than or equal to one (Fudenberg and Tirole, 1991; Bagnoli and Bergstrom, 1989). I demonstrate the logic with a uniform distribution in the Appendix.

[12] As shown in the Appendix, the equilibrium demand x^* is decreasing in certainty until the probability of war reaches zero, then increasing afterward as the minimum possible c_B increases.

opponent—to one of ignorance, where the the range of how strong an opponent might be is itself unknown (Gelman, 2006). Recall that unwarranted certainty refers to an individual's belief that his information is better than it actual is, which produces overprecision in his estimates. I focus on this form of bias, as opposed to alternative forms of irrational behavior, because many studies replicate the finding that overprecision errors are widely prevalent and persistent.

Importantly, there is strong evidence that such errors are not limited to convenience samples—such as a pool of college undergraduates—but are also evident among policy elites who grapple with estimation tasks. Despite the higher stakes in decisions of war, elites and experts are susceptible to making similar errors (Tetlock, 2005; Friedman and Zeckhauser, 2012). For instance, CIA Director Allen Dulles spoke to President Kennedy preceding the Bay of Pigs Invasion, saying "I stood right here at Ike's desk and told him I was certain our Guatemalan operation would succeed, and Mr. President, the prospects for this plan are even better than they were for that one" (Rovner, 2011, p. 25). Carter administration officials were certain that opposition groups during the Iranian Revolution were weak and only a temporary nuisance as mere months before the Shah's fall they agreed to sell arms to the extant regime for force projection in the Persian Gulf.[13] Observers of crisis decision-making are often troubled by the lack of reflection and careful analysis exhibited in the process. When debating using force to assist dissident factions in the Dominican Republic following Trujillo's assassination, Under Secretary of State Bowles was dismayed by others' cavalier assessments about the balance of forces, suggesting "we had no real knowledge of who the dissidents were, their views, or depth of influence . . . and the next order of business was to find out what was going on."[14] Similarly, Michael Forrestal of the National Security Council Staff stated, "I don't have the feeling that everybody is thinking very much" during discussions about intervening in the Laotian Civil War.[15] Overprecision errors occur even when the stakes are enormous.

[13]Policy Review Committee Meeting, 7/5/78, Zbigniew Brzezinski Collection, "PRC 64 7/5/78," Box 25, Jimmy Carter Library.

[14]Memorandum by the Under Secretary of State (Bowles), June 3, 1961, *FRUS*, Volume XII, American Republics, 1961-1963, Document 310.

[15]Memorandum of Telephone Conversation Between the Under Secretary of State for Political Affairs (Harriman)

To model unwarranted certainty, I follow Kurizaki (N.d.) by incorporating error to player A's estimate of B's resolve. Let $\theta \in [0, 1]$ indicate the error in A's perception such that θ represents the percentage uncertainty reduction, in the form of shrinking range boundaries, from the unbiased estimate of B's resolve.[16] For instance, if an appropriate reading of the intelligence on B's cost tolerance indicates that c_B is distributed uniformly between 0.3 and 0.7, then $\theta = 0.5$ indicates A perceives that c_B is distributed uniformly between 0.4 and 0.6.[17] A now optimizes its proposal to B using this new estimate of c_B (0.4 to 0.6) while in fact c_B is drawn from the unbiased distribution (0.3 to 0.7). More concretely, this is akin to A erroneously concluding that B's fatality tolerance is neither extremely high nor extremely low. A optimizes its proposal based on this faulty assessment while in fact B's tolerance could fall anywhere in the full distribution. Thus bounded rationality in the model only applies to an actor's representation of the strategic situation. Conditional on the representation, actor behavior is value maximizing, not satisficing (Bendor et al., 2011).

Players in the model depart from a rationalist benchmark in a limited and well-defined form. Unwarranted certainty is modeled as an actor believing the available information offers more precision about the value of an opponent's resolve than the available information actually provides.[18] This is most easily conceived of as having a prior expectation that is too narrow and potentially excludes the true value of opponent resolve. Selective exposure or perception (see Gerber and Green (1999)) or correlational neglect (Glaeser and Sunstein, 2009; Ortoleva and Snowberg, 2015), are alternative pathways consistent with overprecision, though adjudicating between these alternatives is beyond the scope of this paper.

and Michael V. Forrestal of the National Security Staff, June 9, 1964, *FRUS*, Volume XXVIII, Laos, 1964-1968, Document 85.

[16] In the beta distribution, we can think of θ as reducing variance.

[17] Note that the mean of the overprecise distribution matches that of the correctly assessed distribution. Again, the focus is on the role of certainty and not optimism. Over-optimism's war-promoting effect is intuitive and immediate in this context. The distinction between the two is justified as overprecision does not imply over-optimism. Moreover, there is less consistent evidence of over-optimism than for overprecision because the former is prone to reversals (Moore and Healy, 2008).

[18] This is distinct from greater precision on the estimate of the success parameter in a Bernoulli trial with known and fixed outcome values (Friedman and Zeckhauser, 2015). Rather, in my model the possible outcome values are themselves unknown and precision refers to the distribution of these values.

The relationship of interest is whether increasing the degree of unwarranted certainty affects the ex ante probability of war. Varying θ demonstrates that the probability of war *can* be increasing in unwarranted certainty. The result is particularly pronounced at the extreme where a single point forecast is substituted for an estimate that should have uncertainty. Even when the war probability is not increasing, it decreases more gradually than in the warranted certainty case. This result runs in opposition to the standard rationalist model used for Hypothesis 1a. The upper line in Figure 1 plots the comparative static as θ shifts from 0 to 1 with $c_B \sim U(0.05,0.55)$, $c_A = 0.2$, and $p = 0.4$

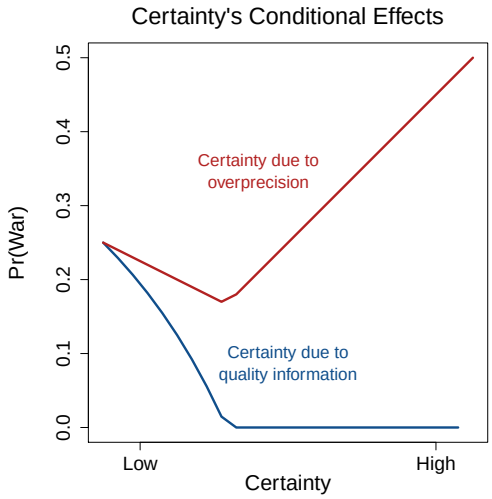

Figure 1: Certainty's conditional effect on war onset. This figure summarizes the analysis. For the lower line, $F(z)$ begins as $c_B \sim U(0.05,0.55)$ and collapses to $c_B = 0.3$. The upper line plots the probability of war for a fixed $F(z)$ as excess certainty (θ) increases. Here $F(z)$ is $c_B \sim U(0.05,0.55)$ and θ goes from 0 to 1.

The non-monotonicity of the relationship requires explanation. The probability of war initially declines because the proposing state A believes the enemy's resolve to be more narrowly prescribed. As a result, the returns to risking war shrink because the perceived difference in the quality of bargains needed to appease a strong enemy versus a weak enemy shrinks. The proposing state diminishes the risk by making less demanding proposals which lowers the probability of war. However, the dynamic changes once A narrows its estimate to the point where it is willing to make

proposals that would satisfy the most resolved enemy (as erroneously perceived by the A). A will make the most onerous demands possible ($min(c_B) + p$) that still satisfy B such that A perceives there is no chance of war. Further false certainty substantively means that A excludes the possibility that B is either extremely resolved or extremely unresolved. As A gets more certain and ignores the possibility that B could be highly resolved, A will make increasingly onerous demands. More onerous demands increase the probability of war. At the extreme case of total excess certainty, A assumes c_B is the mean value of $F(z)$. This produces a 50% probability of war because B's actual resolve is drawn from a distribution with half of its mass on either side of the mean.[19]

To fix concepts, consider the earlier quote from Allen Dulles regarding his certainty about success in the Bay of Pigs operation. A less certain assessment would place some weight on the possibility that Castro and his forces were strong or highly resolved. Granting this possibility would reduce the expected attractiveness of fighting for the US, making it willing to make less onerous bargaining demands, either tacitly or explicitly. However, Dulles precluded the possibility of Castro being a very tough type. Consequently, US demands exceeded anything Castro would accept. Bargaining, either tacit or explicit, fails due to these high US demands which demonstrates how certainty due to overprecision can increase the probability of conflict.

The importance of theorizing certainty's causes should be readily apparent now. The two scenarios are in tension with one another in the absence of a reliable means for establishing the appropriate amount of certainty. The effect of increasing the degree of observed certainty on conflict likelihood is contingent on whether that increase is due to quality information or perceptual error. Understanding the source of the certainty resolves the ambiguous empirical implications.

> *Hypothesis 1b: The ex ante probability of conflict is decreasing at a strictly greater rate in warranted certainty than in unwarranted certainty (where it can increase).*

An implicit assumption in the hypothesis is that there is variation in whether or not estimated certainty is warranted. If all individuals (or groups) in all contexts are equally prone to overpre-

[19]The intuition is similar with a beta distribution, but without the clarity provided by a uniform's shrinking bounds.

cision errors then there is no value to hypothesizing interactive effects. However, there is compelling evidence that individuals systematically vary in their likelihood of approximating rationalist behavior—that is, behavior without overprecision errors (Stanovich and West, 1998; Coelho, 2010). Individual heterogeneity in adherence or divergence from rationalist expectations also occurs among political elites (Tetlock, 2005; Hafner-Burton et al., 2014; Rathbun, 2015). The next section theorizes sources of this heterogeneity as it relates to the proclivity for overprecision.

Some clarifying notes are in order. Can opponents credibly provide information to eliminate errors of overprecision? Within this framework they cannot due to their incentives to misrepresent their strength or type (Fearon, 1995). Opponents would only send a message, either indicating their type or attempting to generate greater uncertainty about their type to correct overprecision errors, when it improves their bargaining position. The estimating state rationally discounts this message given the opponents' incentives.

Why do leaders not adjust their estimates given the potential for overprecision to produce ex post inefficient outcomes? Put simply, individuals typically do not believe they are making errors. Even when warned of overprecision, individuals fail to fully adjust their estimates (Alpert and Raiffa, 1982). Ortoleva and Snowberg (2015) conclude, "it is very difficult to persuade overconfident [overprecise] citizens that their prior is incorrect, as they tend to attribute contradictory information to others' biases." This difficulty to recognize overprecision in one's own estimates also helps explain why there is limited learning within and across executive administrations about this source of error.[20] Similarly, the disinclination to believe one is overly certain distinguishes overprecision from problems of costly information acquisition. Overprecision arises because of a perceptual error regarding the quality and quantity of available information, not from a prohibitively high cost of acquiring additional information.

[20]President Kennedy is a partial exception within the data as there is evidence of more thorough information gathering processes after the Bay of Pigs.

Causes of Certainty

When is certainty most likely to be warranted? I specify an information aggregation process in crises and then theorize circumstances that increase overprecision errors.

Actors and Aggregation

Subjective assessments are by definition the result of individual estimative efforts. Accordingly, leaders and close advisors, as opposed to state or system level variables, serve as the locus of analysis for the theory of certainty's causes. First, these are the actual decision-makers assessing and optimizing foreign policy. Second, regime type is mismatched with the theoretical concepts of interest. In the US, electoral and congressional considerations may affect whether war occurs, but are largely absent from the assessment process itself. Moreover, a regime-centric approach denies actors the agency that is empirically evident. Third, a fruitful burgeoning literature in international security postulates that leaders play a pivotal role in conflict (Chiozza and Goemans, 2011; Saunders, 2011; Horowitz and Stam, 2014). I expand the analysis beyond leaders because executive estimates are intimately shaped by the information and opinions provided by senior advisors. For instance, understanding Henry Kissinger's input is essential to analyzing President Nixon's certainty when deciding to expand US involvement in the Vietnam War into Cambodia.

Understanding how elites contribute to estimated certainty requires a theory of information transmission and aggregation. In hierarchical settings with a principal, such as the president, and an agent, such as an advisor, the credibility of the latter's input is a function of the two actors' preference alignment. With sufficiently aligned preferences, there is an equilibrium in which the agent reports his true perception of the situation and the principal updates his beliefs accordingly (Crawford and Sobel, 1982; Bendor, Glazer and Hammond, 2001). I suggest the president's relationship with advisors approximates these conditions because they share an objective of maximizing the state's payoff in international crises. Krasner (1972) provides a strong defense of this assumption.

Notably, the president selects his own advisors and can sanction them for incompetence, both of which facilitate credible communication (Krehbiel, 1991; Meirowitz, 2006).

This conception differs from the Bureaucratic Politics Model of Allison (1971), in which elites have divergent preferences shaped by parochial interests which undermine credible information transmission. The empirical validity of that Allison model is highly contested (Bendor and Hammond, 1992; Welch, 1992). Context is key for assessing whether bureaucratic position reliably predicts actor preferences. In budgetary matters, the argument for such a link is compelling. However, the evidence is mixed in crisis decision-making settings (despite their potential budgetary implications) with several notable counter-examples, such as Secretary of State Clinton's hawkish position on the Afghanistan surge (Marsh, 2014) or Defense Secretary McNamara's relatively dovish stance during the Cuban Missile Crisis. I model advisors as general welfare maximizers who can credibly communicate with the president.[21]

Due to sufficient preference overlap in crisis settings, I expect that advisors offer their true beliefs to the president and the president rationally views this information as credible. Consequently, aggregated assessments will approximate the average of elite estimates weighted by an elite's access to the executive or floor time in meetings. When advisors with overprecise estimates are highly involved in the deliberation process, the aggregate assessment will be overprecise. Thus estimative errors occurring at the individual level become manifest at the group level.[22]

Note that the focus is on elite advisors with regular access to the president and not on intelligence analysts from lower levels of the bureaucracy. I expect, and observe, that the extent of overprecision is greater among elite advisors than in intelligence reports.[23] The intelligence cy-

[21]Whether elite communication occurs as theorized is testable. If contrary to the supposition, preferences are predictable based on bureaucratic position then cheap talk—positions that match ex ante expectations—will be discounted. Only those messages that go against type, such as a State Department official expressing a preference for using force, will influence presidential assessments. Robustness tests discussed below reject this alternative model.

[22]Overpecision could even be exacerbated at the group level for the same reasons that it occurs at the individual level. Selective exposure or correlational neglect would perpetuate and enhance overprecision errors in a group setting.

[23]This expectation is substantively consistent with the account in Yarhi-Milo (2013). However, this is not to suggest that overprecision is absent from intelligence products. See Friedman and Zeckhauser (2012) for examples.

cle literature (Betts, 1978; Halperin and Clapp, 2006; Rovner, 2011) documents this failure for uncertainty to be properly transmitted and received when passing from analysts to policy-makers, whether due to the latter's scarcity of time or their distrust of the intelligence community.

Sources of Overprecision

What factors reduce the prevalence of overprecision? Put differently, when are actors more likely to recognize the full scope of uncertainty? My hypothesis focuses on the substance of estimation. Recognizing the full scope of uncertainty is more likely when individuals consider all factors that affect the estimated quantity. For instance, overprecision is less likely in estimates of expected war outcomes that consider not just raw military assets, but also an opponent's military training, a net assessment of how opposing sides' forces interact, and an opponent's resolve. Adding dimensions which carry uncertainty to the estimation process counteracts the tendency toward overprecision. Put differently, ignoring salient dimensions and their attendant uncertainty is akin to an overprecision error as doing so produces an overly certain estimate.

Estimation processes frequently marginalize the political, as opposed to military, factors that affect expected conflict outcomes. Consider two ways political attributes shape expected conflict outcomes. First, political qualities such as resolve, unity, and interests directly affect military efficacy. Understanding an opponent's political interests and resolve is essential for estimating conflict's shape, costliness, and outcome. Second, the broader political landscape of an opponent affects post-conflict planning and has ramifications for translating military accomplishments into political gains. Military success is not an end in itself as it must be translated into securing political objectives, such as territory or preferred policy outcomes. The pertinent conflict outcome is not the strictly military result but rather the resulting political division of the contested resource. The importance of linking military and political outcomes has deep roots in the study of war. Clausewitz (1976) famously proposes that war is political to its core and combat strategy must be instrumental rather than an outcome in itself. Schelling (1966) roots war, or the imposition of harm more

broadly, in the context of achieving political aims. In the bargaining model, war is a lottery determining which side secures the political spoils, be it territory, resources, or a preferred slate of policies. However, this is a simplifying abstraction. In anything besides absolute war, and likely even then, battlefield success is correlated with but not equivalent to realizing political goals.

Marginalizing these political factors that affect conflict outcomes causes overprecision by neglecting a source of uncertainty. I theorize that elites who oversee international diplomacy are least likely to neglect these political factors. In the US context, this stipulates that State Department officials are most likely to focus on and be cognizant of the added complications of assessing an opponent's resolve and political landscape. Diplomats differ due to their institutional mandate to oversee foreign policy which requires expertise on an opponent's military *and* political attributes. In contrast, individuals from other bureaucracies tasked with defense and security policy are comparatively more likely to focus on military attributes of conflict and ignore its political dimensions.

Officials from the State Department are less prone to overprecision errors. More generally, when State Department officials are heavily involved in the estimation process, overprecision is less likely. When these officials are included in deliberations, certainty is more likely to stem from quality information. When they are excluded, certainty is more likely to stem from overprecision errors. Linking to the earlier hypothesis, I theorize that the extent of State Department involvement in assessment processes conditions, or moderates, the relationship between certainty and conflict.

Note that the argument is not that State Department officials differ in cognitive dispositions. Borrowing terms from Tetlock (2005) and others, I do not assert that diplomats are foxes while all others are hedgehogs. Nor is the argument about strategic bureaucratic behavior whereby State Department officials inflate their uncertainty to foster diplomacy and help their organization gain prominence or resources (though I address these possibilities and other alternative explanations later). Preference divergence is not the source of estimate variation. Rather, State Department officials are less prone to overprecision because they are institutionally mandated to analyze factors that are important but frequently overlooked. The argument diverges from the "where you stand

depends on where you sit" notion of bureaucratic politics (Allison, 1971, p. 711) and instead contends that *what you consider* depends on where you sit.

To fix ideas I provide examples of policy-makers ignoring salient political factors and then consider examples where State Department officials offered divergent assessments. With reference to post-conflict political planning, President Obama suggested that the US and partners "underestimated" the difficulty of achieving the desired political outcomes in Libya. He continued, "So that's a lesson that I now apply every time I ask the question, 'Should we intervene militarily? Do we have an answer [for] the day after?"'[24] Though startling that such questions are not always asked before conflict, it is indicative of overprecision errors at the highest levels of US decision-making on the most important matters. Similarly, estimating expected outcomes in the Iraq War, at least in hindsight, required estimates of the conventional capabilities and resolve of the Iraqi national forces *and* estimates of the difficulty that US and coalition forces would experience in translating conventional military success into their preferred political outcomes. A prioritization of the former and ignorance of the latter generated overprecise estimates for the expected political outcome of the war. Ignoring political factors also produced excess certainty before the Bay of Pigs invasion. CIA estimates rested on the assumption that Castro had not consolidated power and a quick political uprising would emerge.

State Department analysis frequently casts doubt on the certain projections of others due to its emphasis on the political reasoning and conditions of opponent states. Returning to the Iraq War, officials from the State Department, who were largely excluded from the post-war planning process (Halperin and Clapp, 2006, p. 171) warned military commanders of "serious planning gaps for post-conflict public security and humanitarian assistance ... which would undermine an otherwise successful military campaign, and our reputation internationally."[25] State Department officials even exhibited greater uncertainty on military matters, as its Bureau of Intelligence and

[24]Quotes from Thomas Friedman, "Obama on the World," *New York Times*, August 8, 2014.

[25]Memo to Under Secretary Dobriansky, February 7, 2003, NSA Archive Electronic Briefing Book No. 163.

Research (INR) was the leading skeptic regarding Iraqi weapons programs in the build-up to the 2003 war (*National Intelligence Estimate*, 2002). State Department officials were dubious of CIA estimates about Cuba's political conditions before the Bay of Pigs. However, they were excluded from many deliberations due to CIA's demand for secrecy. Similarly, during the Nixon Administration's discussion of launching incursions into Cambodia, Secretary of State Rogers and his Under Secretaries questioned the estimative certainty held by others and the viability of the new Cambodian regime.

The preceding discussion generates the following hypothesis with two implications.

> *Hypothesis 2: Greater involvement of elites who oversee international diplomacy reduces the likelihood of overprecision.*
>
> - *Implication A:* Given the same estimation task, elites who oversee international diplomacy will form less certain estimates than their counterparts from other organizations.
> - *Implication B:* High involvement by elites who oversee international diplomacy conditions the relationship between certainty and conflict such that certainty causes peace.

Implication A follows immediately. Implication B is the more important expectation as it links the causes and consequences of certainty. The intuition is straightforward. Overprecision is less prevalent among elites tasked with conducting international diplomacy. When these elites are heavily involved, certainty is more likely to be warranted. When certainty is warranted, the standard rationalist expectation holds where certainty facilitates bargaining.

A possible critique of the hypothesis is that who gets the most floor time or access to the executive is not random. Leaders might be inclined in one direction and only seek advice from those suspected to have confirming dispositions. If the president wants a diplomatic solution, he might select to give the State Department more floor time. However, selection of this sort biases results toward finding that greater State Department involvement reduces conflict likelihood but this is not a key implication of the theory.[26]

[26]The theory, roughly, expects that State Department involvement has an indeterminate effect on conflict likelihood

Implication B does not address the direct relationship between State Department involvement and war. Rather, the implication concerns the relationship between certainty and war *conditional* on high State Department involvement. Even conditional on State Department involvement, there is variation in estimated certainty across crises due to different levels of available information. The hypothesis addresses how this variation within instances with high State Department involvement associates with war and how this relationship differs when State Department involvement is low.

Imagine two crises where the president is predisposed toward a peaceful outcome. The selection story suggests the State Department would be highly involved in each estimation process and peace would ensue in both cases. This would not bias results in favor of validating the hypotheses. Due to differing information levels, there will be variation in the estimated certainty in the two crises. I hypothesize that war is more likely in the crisis with less certainty.

In sum, the consequences of certainty are indeterminate without considering whether certainty stems from quality information or overprecision errors. The prominence of diplomats in the estimative process conditions, or moderates, the relationship between certainty and war. Diplomats incorporate uncertainty about opponents' political factors which reduces the prevalence of overprecision. As a result, when decision processes include diplomats, certainty is more likely to reflect quality information and cause peace as predicted in standard bargaining models. When decision processes marginalize diplomats, certainty is more likely to be unwarranted and cause conflict.

4 Research Design

To evaluate the hypotheses I conduct automated content analysis on an original corpus of declassified security documents from instances in which the US considered using force. The rest of this section describes the sampling procedure, construction of a corpus of documents, methods of measuring certainty, and the relevant outcome, explanatory, and control variables.

due to the non-monotonic relationship between unwarranted certainty (θ) and conflict. Rather than observing an indeterminate relationship, selection bias causes high State Department involvement to reduce conflict likelihood.

Opportunities to Use Force

An opportunity to use force is the unit of analysis. Existing definitions for opportunities to use force are insufficiently strict for my purposes (Fordham, 1998; Meernik, 2004). I define an opportunity to use force as any instance in which a state's leader and senior advisors seriously discussed using its own military to strike abroad. This definition excludes analyses that occur at lower levels in bureaucracies (such as Department of Defense war-gaming), minor shows of force with no intention or discussion of actually using it, and decisions restricted to whether to arm or fund combatants in war. The objective is to identify instances where decision-makers faced a clear choice between costly conflict and peaceful measures.

The sample is restricted to US opportunities to use force during the Eisenhower through Carter administrations. There are benefits to the sampling restriction as it holds many potential confounders constant. First, keeping the assessing state constant mitigates concerns that institutional differences across states are driving the results. As one of two superpowers, US behavior is more than a parochial concern, but rather of general interest. Second, background characteristics that may affect estimated and expressed uncertainty—such as gender, education, and race—were virtually constant among elites. Third, the strategic environment of the international system during the Cold War was relatively similar, albeit with periods of detente, across the administrations. Fourth, on a practical note, these temporal bounds also reflect the availability and consistency of the main documentary evidence source. However, these spatial and temporal bounds raise external validity concerns which I address later.

The sample is primarily drawn from two sources. First, I collect crises involving the US as an actor as defined by the International Crisis Behavior (ICB) dataset (Brecher and Wilkenfeld, 1998). Second, borrowing from Fordham (1998), I include instances with major military mobilizations, which are coded as level one and two uses of force by Blechman and Kaplan (1978, p. 50). If many military assets are directed toward a possible conflict location, this is a good indication of serious discussions about using force. These two sources contribute 46 observations. I supplement

the sample with 15 additional observations. These include some of the more critical level three and four military mobilizations coded by Blechman and Kaplan (1978), militarized interstate disputes (not all of which included the use of force) (Ghosn, Palmer and Bremer, 2004), and debates concerning preventive strikes against the nascent nuclear programs of the Soviet Union and China. I present robustness results excluding these additional 15 cases throughout the results discussion.[27]

The criteria yield a sample of 61 observations, listed in Table 1. Some observations concern potential conflict onset—such as the 1954 fall of Dien Bien Phu, 1958 turbulence in Lebanon, 1969 North Korean EC-121 spy plane incident, and 1975 Cambodian *Mayaguez* seizure. Others capture pivotal moments to potentially expand the scope of ongoing conflicts—such as 1953 deliberations surrounding the Korean War armistice and 1971 invasion of Laos during the Vietnam War. There is wide variation in the explanatory variables across the observations. State Department involvement in assessments varies from instances of close collaboration between Eisenhower and Dulles to cases of high marginalization during the Nixon administration. Additionally, the extent of available information varied from cases with poor intelligence (such as the Iranian Revolution) to cases with relatively rich intelligence (such as the Soviet invasion of Czechoslovakia).

Corpus

To measure certainty I collect documentary evidence from elite deliberations about using force. Focusing on elites provides a direct measure of uncertainty in crisis bargaining. Declassified reports made available through the Foreign Relations of the United States (FRUS) series are the core sources for these estimates. FRUS is the official documentary record of the State Department, though it also includes sources from Presidential libraries, Department of Defense, National Secu-

[27]See the Appendix for more on sampling criteria. Sample censoring by excluding cases where using force never received scrutiny could introduce bias if the cause of the censoring relates to the level of certainty. To bias results in favor of the hypothesis for cases where I anticipate certainty is warranted, the cases excluded due to sample censoring, which definitionally ended peacefully, would need to have low certainty estimates. The inverse would need to hold for cases where I expect unwarranted certainty. Thus, to confirm the hypothesized moderated relationship between certainty and conflict, the censoring bias would have to flip directions depending on State Department involvement.

Table 1: Opportunities to Use Force

Year	Observation	Year	Observation
1953	Chinese Offshore	1964	Laos I
1953	Korean War Armistice	1964	Panama Flag
1953	Soviet Nuclear*	1964	Tonkin Gulf
1954	Dien Bien Phu	1965	Dominican Republic
1954	Guatemala	1965	Pleiku
1954	Taiwain Straits I	1967	Six Day War
1956	Suez	1968	Czechoslovakia
1957	Jordan I	1968	Pueblo
1957	Syria	1968	Tet
1957	Taiwan Straits Turmoil*	1969	Cambodia Bombing
1958	Berlin Deadline	1969	Korea EC-121
1958	Indonesia*	1970	Black September
1958	Jordan II	1970	Cambodia Invasion
1958	Lebanon	1970	Cienfuegos
1958	Taiwan Straits II	1971	Cuban Shipping*
1960	Cuba*	1971	Laos II
1961	Bay Of Pigs	1972	Christmas Bombing
1961	Berlin Wall	1972	Ports Mining
1961	Kuwait*	1973	Arab Israeli War
1961	Pathet Lao	1973	Libya*
1961	Phuoc Vinh	1974	Cyprus II*
1961	Trujillo	1975	Angola
1962	China Arms Buildup*	1975	Mayaguez
1962	Cuban Missile Crisis	1975	Saigon Fall
1962	Nam Tha	1977	Uganda
1962	Taiwan Return*	1978	Iran Revolution*
1963	China Nuclear*	1978	Shaba II
1963	Haiti*	1979	Afghanistan
1963	Yemen*	1979	Cuban Soviet Brigade*
1964	Congo	1979	Iran Hostage
1964	Cyprus I		

* indicates qualification with less restrictive sampling criteria

rity Council, Central Intelligence Agency, and other agencies and individuals involved in crafting US foreign policy. Documents include National Security Council meeting transcripts, intra-elite memos (such as memos from Kissinger to President Nixon), and minutes from elite conversa-

tions in both unstructured (such as ad hoc White House meetings) and structured settings (such as presidential briefings from the Joint Chiefs of Staff).[28] Content from multi-speaker transcripts is separated into speaker-specific texts to gauge the effect of organizational role on estimative certainty. Due to the limited availability of FRUS volumes covering the Carter administration, I attempted to gather the equivalent documentary evidence from the Carter Presidential Library archives. Transcripts and memos akin to those found in FRUS were widely available for the six Carter-era observations.[29]

FRUS primarily consists of private memos and transcripts. Public statements are subject to strategic considerations which alter the expressed certainty. Private conversations are more likely to reflect sincere beliefs. When opting to intervene in the Laotian Civil War in 1964, Johnson told advisors, "we should go ahead with the mission but that he had doubts about the action."[30] The president making this statement publicly is nearly inconceivable.

FRUS volumes are typically organized by region and time period and consequently volumes contain documents that are unrelated to the observation of interest. Even when related to the observation, many documents contain nothing about the use of force and are inappropriate for measuring certainty over expected conflict outcomes. Accordingly, I read and prune the set of documents, retaining only those directly addressing issues related to the theory. For expositional clarity in the model, uncertainty only pertained to resolve. For the empirics I want to capture uncertainty over any factor affecting the bargaining range. This includes resolve, other salient

[28]Data censoring within documents is of limited concern. Classified segments within available texts typically concern targeting location details, such as in North Korea. This is only a minor inconvenience because I am concerned with how leaders address uncertainty about conflict generally, not how they assess specific targets. Whether entire documents remain classified is a more difficult problem to assess. That said, the availability of documents with estimation errors (judged ex post) suggests that ongoing classification is unlikely to correlate with estimation errors in a way that biases results.

[29]Partial exceptions include a limited record of input from the Joint Chiefs of Staff and little evidence from deliberations in the two weeks preceding the April 1980 Iran Hostage rescue attempt. Additionally, I supplement FRUS documents for the Cuban Missile Crisis with transcripts from the Executive Committee of the National Security Council (ExComm).

[30]Memorandum of Conference with President Johnson, June 8, 1964, *FRUS*, Volume XXVIII, Laos, 1964-1968, Document 83.

political attributes, and military capabilities. The theoretical implications of using these factors are equivalent to using only resolve. I restrict the corpus to memos sent to or by the president and transcripts from meetings which the president attended. Given his primacy in the decision-making process, aggregate estimates should only be informed by opinions to which he was privy.[31]

In total, the corpus encompasses 1,109 speaker texts generated from 364 documents. 85% of speaker texts come from meeting transcripts with the remainder from presidential memos. Across the full document set there are 176 unique speakers and over 132,000 total words.

Measuring Uncertainty

Testing the hypotheses requires measurements of uncertainty at the level of each speaker, bureaucracy, and observation. I measure uncertainty with computer-assisted text analysis using a dictionary method similar to that found in McManus (forthcoming). The analysis uses the "If" dictionary, from the Harvard General Inquirer, constructed to measure uncertain tones. The dictionary includes words such as "approximately," "doubt," "possibility," and "unpredictable."[32]

Using the dictionary, I calculate the relative frequency of uncertain words as a percentage of total words in each text. Following prior work summarized in Grimmer and Stewart (2013), I construct measures by assigning a score of one to uncertain words, summing across the document, and dividing by the total number of words in the document. This basic procedure can provide measures of uncertainty at multiple levels. First, I split texts by speaker and construct a speaker level measure of uncertainty for each observation. Second, I aggregate speakers by bureaucratic role—such as including the content of a Secretary as well as an Under Secretary in creating a State Department uncertainty measure for each observation. Third, I construct an aggregate observation level uncertainty score. The first two measures help validate hypotheses about the causes of over-

[31]The 1976 North Korean Tree Trimming Incident meets the sampling criteria but FRUS does not contain documents with the president's involvement so it is excluded.

[32]Dictionary is available at `http://www.wjh.harvard.edu/~inquirer/`. I make small edits to tailor the set of words to the context of theoretical interest. For instance, I add "risk" to the dictionary because decision-makers often voice uncertainty by invoking risks.

precision. The third measure, interacted with whether causes of overprecision are present, is used to test the consequences of certainty. Figure 2 plots the uncertainty score values for some salient observations across the distribution of uncertainty values.

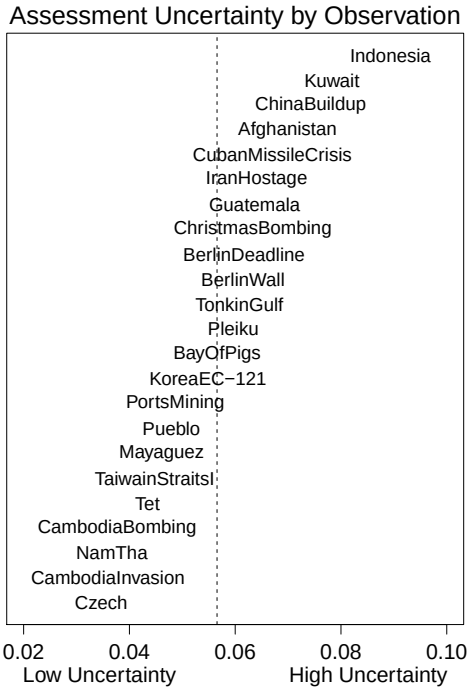

Figure 2: Uncertainty score for some observations. Vertical line plots the mean score.

Dictionary methods are difficult to validate (Grimmer and Stewart, 2013). As elaborated upon in the Appendix, I adopt three approaches for assessing construct validity. First, I analyze instances where there are clear expectations for the relative degree of uncertainty. For example, I examine texts that refer to another text as being too certain. During the EC-121 incident Defense Secretary Laird sent a memo to President Nixon suggesting that the Joint Chiefs of Staff (JCS) estimate of what outcome US forces could impose on North Korea is overly certain. Laird cites risks and doubts about the efficacy of using force. Some text includes:

If U.S. losses occur in the strike (and I believe there is more chance they may than

> the JCS papers indicate)... It is not clear we have the capability now to handle a major confrontation in Korea, if the North Koreans should react with a major assault of any duration against South Korea.[33]

Laird's memo should register a higher uncertainty score than observed in JCS communications. Indeed this is the case as Laird's memo has an uncertainty score of 6.2% compared to the JCS memo with an uncertainty score of 2.6%. For reference, speaker texts with at least 50 words have a mean uncertainty score of 5.2% and standard deviation of 2.6%.

A second approach hand codes a subset of documents (roughly 5%) from the corpus into three categories or degrees of uncertainty. The third approach uses the same subset of documents but assigns uncertainty scores using a Bradley-Terry model based on pairwise comparisons of the relative uncertainty in two randomly selected documents from the subset (Bradley and Terry, 1952; Turner and Firth, 2012). Results from both approaches are positively correlated (53% and 34% respectively) with the dictionary method uncertainty scores.

Outcome, Explanatory, and Control Variables

Conflict is the main outcome variable. It is a dichotomous indicator for the use of force by the US. I exclude threats or demonstrations of force as these may be bargaining tactics rather than evidence of bargaining failures. The outcome coding corresponds to whether the US actions were costly in that there was added destruction to either US military assets or caused by US military assets due to behavior ordered or authorized by the president. Providing arms or assistance to an aligned force does not qualify (as during the Six Day War) unless there is overt US military action or oversight (as in Indonesia in 1958 and the Bay of Pigs). For intra-war observations with the possibility of escalation, the mere response to an attack does not qualify—such as rolling back the Tet Offensive. Rather, the variable is only coded one when there is a significant expansion of the

[33]Memorandum from Secretary of Defense Laird to President Nixon, April 18, 1969, *FRUS*, Volume XIX, Part 1, Korea, 1969-1972, Document 17.

extant operation—such as mining North Vietnamese harbors.

Certainty is the main explanatory variable. It reflects the percentage of total words in each observation that express uncertainty subtracted from the maximum uncertainty score. A measure of certainty, as opposed to uncertainty, eases interpretation of the results. To test the anticipated conditional relationship between certainty and conflict, I use an indicator variable for the magnitude of State Department involvement. To reflect the State Department's role, I calculate the share of total words per observation that come from State Department officials, which varies from 0% to over 50%. *State Involvement* is coded one for the observations above the median value and zero otherwise. Recall that the extent of State Department involvement proxies for the presence of overprecision errors and conditions the relationship between certainty and conflict.

I control for variables that might confound the relationship between certainty and conflict—that is, variables that are a common cause of the explanatory and outcome variables. Summary statistics for all variables are in the Appendix.

- *Relative Capabilities:* The balance of capabilities between the US and the target likely affects both certainty and conflict. Dyads near parity are assumed to have relatively high degrees of asymmetric information and shown to be more conflict prone (Reed, 2003; Bennett and Stam, 2004). The variable uses CINC scores to measure US capabilities over the sum of US and target capabilities (Singer, Bremer and Stuckey, 1972).

- *Regime Type:* A target state's regime type can influence uncertainty through multiple pathways. For instance, if media transparency correlates with regime type and transparency reduces uncertainty, then there will be less uncertainty when confronting more liberal regimes. The vast democratic peace literature (Oneal, Russett and Berbaum, 2003) finds that target regime type affects US willingness to use force. I use Polity IV scores (Marshall and Jaggers, 2002) to measure regime type.

- *Proximity:* Geographic proximity increases the likelihood of conflict. It might plausibly

reduce uncertainty if greater distances impede intelligence gathering. I use distance in kilometers between capital cities for the measure (Gleditsch and War, 2001).

- *Party:* Holding an extreme political ideology correlates with overprecision, with the relationship being more pronounced for conservatives (Ortoleva and Snowberg, 2015). The variable is an indicator equal to one for Republican presidents and zero otherwise.

- *Age:* Older individuals might be more prone to unwarranted certainty (Ortoleva and Snowberg, 2015). Additionally, older leaders in democracies are more likely to engage in conflict (Horowitz, McDermott and Stam, 2005). The variable measures the president's age.

- *Non-state Enemy:* Some observations in the data set are opportunities to use force against non-state actors. For instance, threats to Jordan's King Hussein during the 1950s led the US to consider using force. The US may have less intelligence about non-state actors and a different disposition toward conflict with them. I control for this with a binary variable.

- *Combat* captures a president's exposure to combat with Nixon and Carter equal to zero and all others equal to one. Leaders with combat experience are less likely to use force than those who served but never saw combat (Horowitz and Stam, 2014).

5 Results

I begin the results presentation with descriptive plots of the data before turning to regressions. Figure 3 plots the key explanatory and outcome variables. The left panel plots the relationship between certainty and conflict, showing a weak association in the direction expected by typical bargaining models. The right panel differentiates observations based on the level of State Department involvement. Greater certainty in observations with high State involvement (blue circles) is clearly associated with peace. This relationship is muted and slightly reversed for observations

with limited State involvement (red diamonds). Relationships observed in the data before making parametric assumptions accord with the theory.

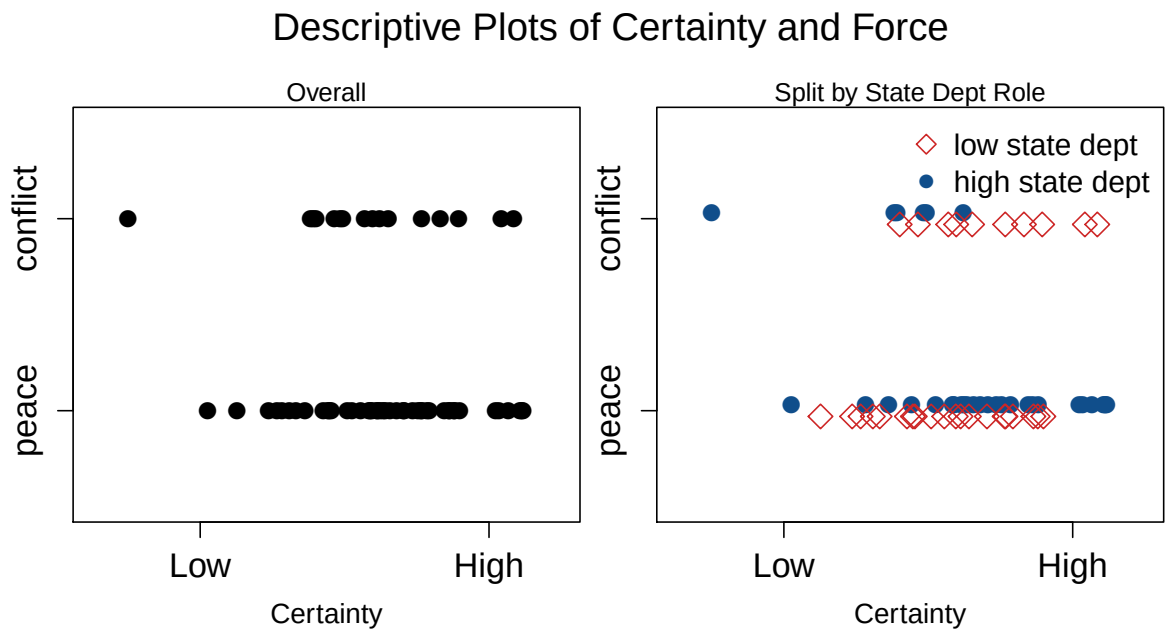

Figure 3: Descriptive plots of the data. Left panel shows the unconditional relationship between certainty and conflict. Right panel differentiates by State Department involvement.

I now ensure the relationships hold when controlling for likely confounders. The first test evaluates the standard rationalist prediction, stated in Hypothesis 1a, that greater certainty decreases conflict probability. After showing limited support for this contention I evaluate State Department involvement. This analysis demonstrates that State officials express greater uncertainty and that their involvement moderates the certainty-conflict relationship in accordance with Hypothesis 2.

Rationalist bargaining theories of war predict a monotonic relationship with the probability of conflict decreasing in certainty. Model 1 of Table 2 tests this prediction with a bivariate regression of conflict on the certainty measure. Due to the binary outcome variable I use logistic regression. While the coefficient on certainty is in the predicted direction, it falls short of conventional levels of statistical significance. That is, greater certainty is only weakly associated with peace.

Controlling for potential confounding variables in Model 2 strengthens the substantive and statistical relationship between certainty and conflict ($p = 0.09$). Figure 4 plots the relationship between the variables. As theorized by the lower line in Figure 1, certainty corresponds with peace. In simulations shifting the certainty explanatory variable from its 25th to 75th percentile values, conflict probability drops from a baseline rate of 24% by 11% points but with 95% confidence bounds spanning from a -32% to 1% point change.

As noted earlier, 15 observations qualify for the sample through more ad hoc criteria. I drop these observations to assess robustness to various sampling procedures. Doing so attenuates the baseline rationalist prediction linking certainty to peace as shown in Model 3. Another potential limitation of the data is that the volume of documentary evidence varies across observations. Unsurprisingly, there is more text for the Cuban Missile Crisis than for the Jordanian troubles in 1957. The total word count for observations varies between 68 and 12,306. To ensure that results are not dependent on observations with limited available evidence, I drop all cases with fewer than 500 total words which reduces the sample to 51 observations. As shown in Model 4, *Certainty* does not reach statistical significance in this specification. In sum, the evidence is directionally consistent with predictions from the traditional rationalist framework but often does not reach conventional levels of statistical significance.

State Department Involvement as a Moderator

I now consider how high State Department involvement affects estimated uncertainty and moderates the relationship of interest. Implication A of Hypothesis 2 is that given the same estimation task, State Department officials, less prone to overprecision, will estimate more uncertainty than elites from other branches. To test this implication I construct observation level uncertainty scores split by institutional affiliation. Each speaker is assigned one of six bureaucratic roles.[34]

[34]Some speakers from departments that rarely appear in security discussions—such as the Treasury Department—are excluded from this analysis.

Table 2: Certainty and Conflict

	(1)	(2)	(3)	(4)
Certainty	-0.20	-0.62*	-0.74	-0.66
	(0.26)	(0.36)	(0.52)	(0.48)
State Dept Involvement		-1.82**	-2.06**	-1.69*
		(0.89)	(1.04)	(0.88)
Combat		-0.53	0.11	-0.30
		(0.95)	(1.07)	(0.98)
Relative Capabilities		16.75	15.32	14.68
		(11.48)	(12.32)	(9.79)
Regime Type		0.23**	0.32*	0.20
		(0.12)	(0.16)	(0.12)
Proximity		0.00*	0.00	0.00
		(0.00)	(0.00)	(0.00)
Republican		-0.08	0.43	0.30
		(1.13)	(1.30)	(1.18)
Age		-0.02	-0.05	-0.05
		(0.08)	(0.09)	(0.09)
Non-State Enemy		-1.08	-1.62	-1.00
		(1.29)	(1.49)	(1.40)
Constant	-0.16	-12.12	-7.97	-8.16
	(1.19)	(11.66)	(13.32)	(9.88)
N	61	61	46	51

$^{*}p < 0.1$ $^{**}p < 0.05$ $^{***}p < 0.01$

Notes: Logistic regression with opportunities to use force as unit of analysis. Models 1 and 2 include the full sample of observations. Model 3 includes only those observations in ICB or meeting Blechman and Kaplan's levels one and two uses of force. Model 4 excludes observations with fewer than 500 words of text.

These categories are President, State Department, Defense Department, CIA, JCS, and White House advisors which consists of National Security Council staff and other advisors. I account for observation-specific factors by de-meaning each bureaucracy's uncertainty score by that observation's aggregate uncertainty score. This within-observation approach yields cleaner comparisons than a between-observation approach that requires controlling for all confounders.

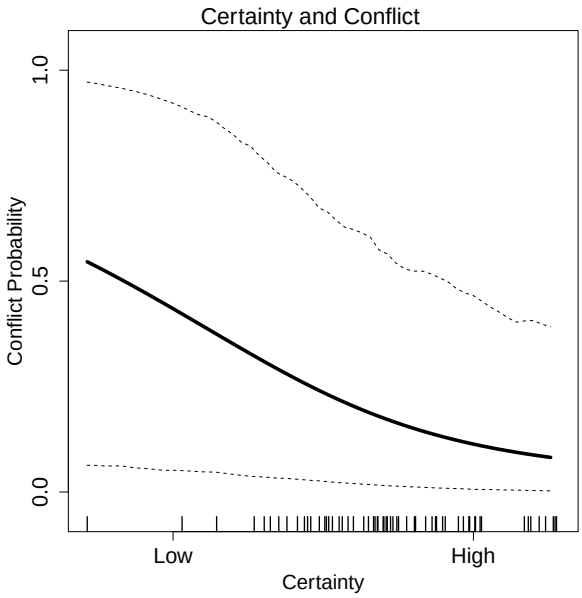

Figure 4: Predicted probabilities of conflict across the range of certainty values with 95% confidence intervals. Based on Model 2. Rug plot reflects data distribution across the explanatory variable.

The results presented in Figure 5 are consistent with the hypothesis. Boxplots show the range of de-meaned uncertainty values for each institution across the observations.[35] The aggregated State Department estimate tends to be more uncertain than estimates from other institutions. Regressing the de-meaned uncertainty scores on a dummy for State Department confirms this relationship ($p < 0.01$).[36] Conditional on crisis specific factors, State officials estimate greater uncertainty.

Does high State Department involvement condition, or moderate, the relationship between certainty and conflict? As stated in Implication B, I expect certainty decreases conflict likelihood when the State Department is heavily involved and increases (or is approximately flat) otherwise. Models 5 and 6 of Table 3 test this contention by interacting the State Department variable with certainty score, without and with controls respectively. In both specifications the coefficient on

[35]I exclude instances where an institution had less than 200 words of total text for an observation. Results are substantively similar without this restriction.

[36]This relationship holds ($p < 0.02$) when excluding all observations from the Joint Chiefs of Staff, which are more certain on average.

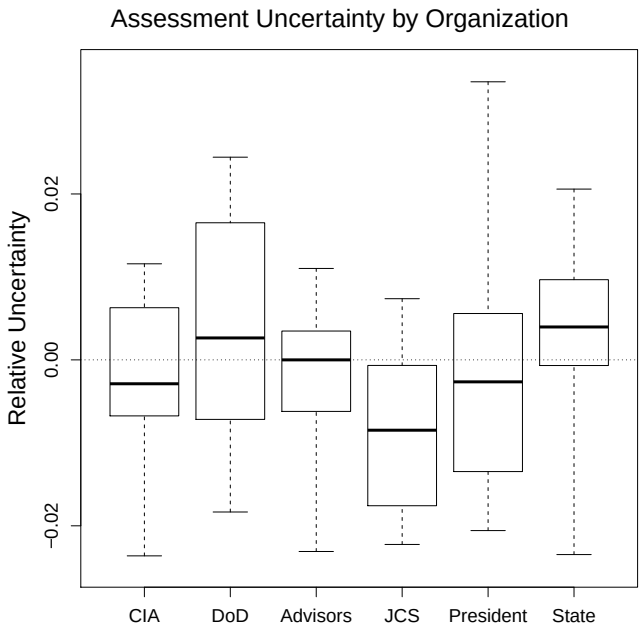

Figure 5: Uncertainty by institution. Boxplot shows uncertainty by institution adjusting for crisis level factors. Bold line indicates median, box length spans from the first to third quartile values, and whiskers extend to minimum and maximum values. The Advisors category encompasses White House staff and dedicated staff members of the National Security Council.

the interaction term is negative and statistically significant. Substantively this suggests that the relationship between certainty and conflict differs depending on State Department involvement. This accords with the theory which states that the slopes of the association between certainty and conflict should differ depending on whether certainty stems from quality information, as proxied by State Department involvement. A likelihood ratio test between Models 2 and 6 confirms that including the interaction improves the model's fit ($p < 0.05$).

I simulate substantive quantities from Model 6 for cases with and without high State Department involvement to facilitate interpretation. Figure 6 plots the results. When State is central to the estimative process—such as when Johnson addressed the *Pueblo* incident—certainty reduces conflict likelihood. Shifting the certainty variable from the 25th to 75th percentile in such cases reduces the conflict probability by 17% points (with 95% confidence bounds of -54%,-1%) from

Table 3: State Department: Moderating Certainty and Conflict

	(5)	(6)	(7)	(8)
Certainty	0.84*	0.33	0.07	0.46
	(0.48)	(0.60)	(0.64)	(0.68)
State Dept Involvement	7.86**	5.69	12.65*	14.09**
	(3.17)	(3.84)	(7.31)	(6.68)
Certainty*State Dept	-1.96***	-1.73**	-3.32**	-3.60**
	(0.71)	(0.87)	(1.68)	(1.55)
Combat		-0.74	0.29	0.16
		(1.07)	(1.41)	(1.36)
Relative Capabilities		19.72	17.96	15.74
		(20.76)	(19.77)	(14.50)
Regime Type		0.26**	0.36*	0.30
		(0.13)	(0.19)	(0.19)
Proximity		0.00	0.00	0.00
		(0.00)	(0.00)	(0.00)
Republican		-0.17	0.75	1.04
		(1.24)	(1.58)	(1.53)
Age		-0.01	-0.14	-0.12
		(0.08)	(0.11)	(0.11)
Non-State Enemy		-0.76	-2.00	-1.27
		(1.39)	(1.72)	(1.83)
Constant	-4.55**	-19.00	-9.55	-10.18
	(2.27)	(21.37)	(20.76)	(15.75)
N	61	61	46	51

$^{*}p < 0.1$ $^{**}p < 0.05$ $^{***}p < 0.01$

Notes: Logistic regression with opportunities to use force as unit of analysis. Models 5 and 6 include the full sample of observations. Model 7 includes only those observations in ICB or meeting Blechman and Kaplan's levels one and two uses of force. Model 8 excludes observations with fewer than 500 words of text.

a baseline probability of 24%. That is, increasing certainty when the State Department is heavily involved is associated with a nearly 75% relative decline in the likelihood of conflict. This corroborates the prediction that when certainty is more likely to be warranted, or based on quality information, it will be associated with peace. In contrast, in the half of the sample with low

State involvement—such as during Bay of Pigs deliberations—there is likely overprecision and the certainty-conflict relationship flips directions. This finding is robust to using alternative sampling criteria as shown in Models 7 and 8.

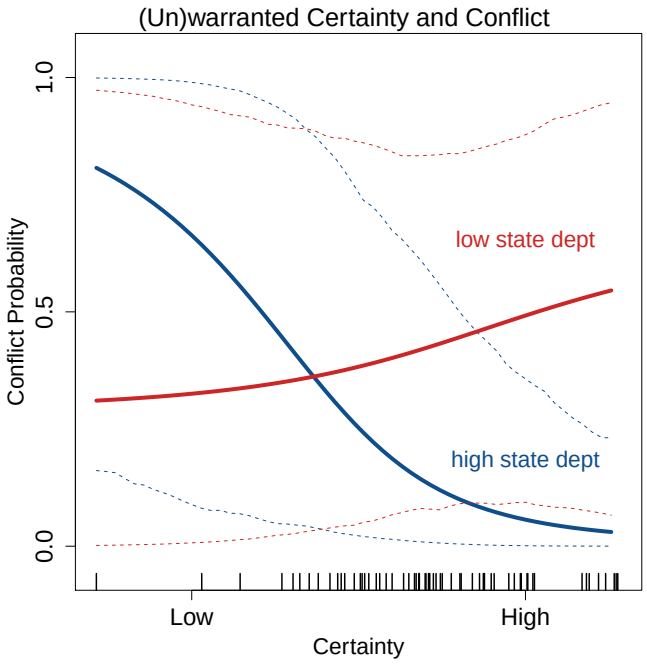

Figure 6: Predicted probabilities of conflict across the range of certainty values with 95% confidence intervals. Based on Model 6 where State Department involvement moderates the relationship. Rug plot shows data distribution.

Recall that results will not be biased in favor of finding an interactive effect even if the president is predisposed toward a peaceful outcome and thus selects to have high State Department involvement. The primary effect of interest is not whether high State Department involvement is associated with peace, which likely suffers from selection bias. Instead, I assess whether certainty is associated with peace conditional on high State Department involvement. Consider two observations for concreteness. In the 1962 Nam Tha and 1965 Dominican Republic crises, Kennedy and Johnson respectively may each have initially preferred peaceful outcomes. They each sought high State Department involvement. However, due to differing information environments, the aggregated estimate was far more certain in the Nam Tha case than in the Dominican Republic case

(respective uncertainty scores of 3.7% and 5.4%). Consistent with my theory for when the State Department is highly involved, conflict occurred in the crisis with greater uncertainty.

The results strongly accord with the hypothesis that certainty's sources moderate the relationship between certainty and conflict. High State Department involvement affects estimated certainty in a manner consistent with there being less overprecision. In such cases, certainty is associated with peace as predicted. When State involvement is low, the sign of the relationship flips.

6 Robustness, Alternative Explanations, and External Validity

This section addresses robustness specifications (results shown in the Appendix) and then speaks to alternative explanations that might be consistent with the findings. Initial checks examine alternative specifications and control variables. One model controls for *Experience*, which reflects the president's time in office. Lessons from prior crises may affect subsequent crisis decision-making (Potter, 2007). However, I find no evidence for this contention and all results for the core explanatory variables remain unchanged. Additionally, I interact *Experience* with *Certainty* and find no evidence that prior experience moderates the relationship between certainty and conflict. This is consistent with leaders struggling to recognize and correct overprecision errors. Perhaps learning or changing dispositions toward assessment certainty and using force occurs at the national level, rather than the individual level. For instance, a war weary public could limit the viability of using force in subsequent disputes. However, incorporating a time trend or a *Recent War* variable measuring years since the last costly US war (Korea or Vietnam) does not substantively change the results. Results are similarly unchanged when including control variables for the number of unique speakers in an observation, whether the observation was initiated by an attack against US forces, or the observation included a compellent threat (Sechser, 2011). Results are also substantively and statistically similar when dropping leverage points or observations from various crises within the Vietnam War, which constitute 8 of the 61 observations.

Other robustness tests use alternative standard error specifications. These include clustering errors on the presidential administration or the opponent as well as calculating bootstrapped standard errors with and without clustering. The central result showing the State Department's moderating effect remains statistically significant at conventional levels across all models.

Two additional checks support important assumptions. First, there is no meaningful difference in either the certainty scores across the two document types—memos versus transcripts—or the mix of document types for State Department speaker texts versus texts from other branches. This assures that higher State Department uncertainty is due to substantive differences, rather than mechanical ones. Second, I examine the most important speakers in each administration measured by amount of text. An assumption used for the information aggregation process is that the volume of text provided by a speaker or bureaucracy is representative of its importance for the final aggregated assessment. A concern is that the president's most important advisors may need few words in order to meaningfully impact the president's final estimate of the situation. As shown in an Appendix table, the assumption appears justified as the corpus mirrors the well-known prominence of key advisors, such as John Foster Dulles, Kissinger, and Brzezinski.

Small Sample Size

The relatively small sample of observations could impede estimation and inferences. Though there are few rules for what constitutes a sufficient sample size, some suggest Type II errors, or false negatives, are particularly likely when using maximum likelihood estimation with less than 30 observations per independent variable (Hart and Clark, 1999). Small samples can also bias logit estimates away from zero (Firth, 1993; McCaskey and Rainey, N.d.). I address these concerns in three ways. First, I estimate all models using least squares because it requires less from the data than logit models. Second, I use Firth's (1993) penalized maximum likelihood estimate which reduces both bias and variance. Results for the State Department moderation effect remain substantively and statistically significant across both tests, while there is little support for

the baseline rationalist hypothesis.

Third, I use randomization inference which does not require an appeal to large samples or parametric models as randomization itself provides a "reasoned basis for causal inference" (Fisher, 1935; Rosenbaum, 2002; Glynn and Ichino, forthcoming). The intuition is that the observed data represents one realization of randomizations of the explanatory variable values across the observations. If a sharp null effect holds, the observed relationship between the realized explanatory variable values and the outcome variables should not significantly differ from the relationship observed in alternative randomization realizations. Put differently, permuting explanatory variable values across observations while holding outcomes fixed should produce substantively similar relationships between the variables to the one observed in the actual data if the null effect holds.

I run 5,000 randomizations for each of the main models (Models 2 and 6) and calculate a test statistic for each. Figure 7 plots the distribution of these test statistics. The left edge of the fill under the density curve indicates the test statistic from the observed data. Focusing on the State Department moderating effect (right panel), only 3% of realizations have a test statistic more extreme than the one found in the data. Results are less compelling for the baseline rationalist model.[37] In sum, State Department involvement conditions the relationship between certainty and conflict across many robustness tests.

State Department Officials: Marginalization and Assessments

It is substantively important to understand when the State Department assumes a limited role. Although this process does not bias results in favor of the hypothesis as discussed above, it could have implications for the generalizability of the results. For instance, if the background covariates systematically differ between high and low State Department observations, then the relationship between certainty and conflict in the latter case may not be applicable to the former. There is some systematic variation as the State Department is more likely to be highly involved in assessment

[37]The Appendix contains a more detailed account of the procedure for calculating the test statistic.

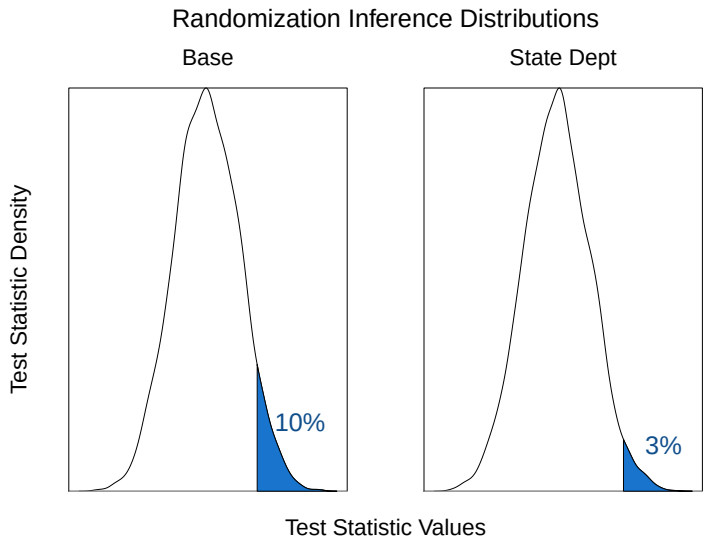

Figure 7: Randomization inference results. Portions of density plots with blue fill represent the percentage of randomizations that produce a test statistic more extreme than those observed in the actual data.

processes when the balance of capabilities is more favorable, the opponent is less autocratic, or the opponent is a non-state actor. However, the central result holds after accounting for these differences either through pruning the data—such as excluding cases involving the USSR to improve balance on relative capabilities—or through propensity score matching (Sekhon, 2011). There is little evidence that State Department officials are systematically excluded when crises are highly militarized. For instance, compellent threats and crisis-initiating attacks against US forces have no statistically significant effect on whether State Department officials are highly involved. Though some factors affect the role of State Department officials, relevant counterfactuals do exist and the core finding of a conditional relationship between certainty and conflict persists within a sample of more comparable observations.

I theorize that State Department officials have greater uncertainty due to their attention to and recognition of the complexity of estimating opponents' political attributes. However, there are plausible alternative explanations for their elevated uncertainty. Individuals who select into the

State Department may systematically differ from those in other bureaucracies. There is evidence that they are more liberal (Milner and Tingley, 2015) and more likely to be generalized trusters (Rathbun, 2012). If either characteristic is associated with less cognitive closure or ambiguity aversion, it could account for State Department officials expressing greater uncertainty (Jost et al., 2007). More generally, there is evidence that elites vary in their cognitive dispositions (for example, Hermann (1980); Hafner-Burton et al. (2014)). If the case, pre-existing dispositions rather than job-specific factors cause higher uncertainty. Empirically assessing this alternative is difficult but one crude test is to examine assessments by officials who held multiple roles. Kissinger is the most notable example, having served as the National Security Advisor (NSA) before becoming Secretary of State. Conditional on observation factors, Kissinger exhibited greater uncertainty when with the State Department, going from being more certain than the average while NSA to less certain when with State. Though subject to confounding, such as having gained experience while NSA, this is at least consistent with my argument as opposed to a pure selection account.

Another possibility is that State Department officials strategically invoke greater uncertainty. Doing so may reflect skepticism about military force's efficacy, belief in diplomacy's ability to persuade other states, or a debating tactic to marshall resources toward diplomatic efforts. Strategic communication of this sort raises the prospect of cheap talk whereby messages by State Department officials expressing high uncertainty are heavily discounted when choosing whether to use force. If uncertainty is strategically used to encourage peaceful outcomes, then instances where State Department officials express greater certainty than others—that is, they go against type—should have a high probability of conflict. However, the empirical relationship between the relative certainty of State Department officials and conflict is weak. Substantively, diplomats having a certainty score 1% point greater than others increases the probability of conflict by 5% points, but with 95% confidence bounds of -7% and 17%.

Pessimism, Uncertainty, and Bureaucratic Politics

Another potential critique is that uncertainty is really a proxy for pessimism and State Department officials are simply more pessimistic about using force. Using a carefully pruned version of the "Ngtv" dictionary from General Inquirer, I construct a pessimism score for each text in the identical fashion to the uncertainty variable.[38] Doing so addresses multiple concerns. First, the uncertainty measure is not a proxy for pessimism, at least as reflected in these dictionaries. The correlation across all speaker texts is -0.05 and across all observations is -0.11. That is, greater uncertainty correlates with slightly *less* pessimism. Second, State Department officials are not systematically more pessimistic in their assessments. De-meaned for observation-specific factors, the average State Department estimate is statistically indistinguishable from the overall estimate. Officials from State have greater uncertainty, not greater pessimism. Third, including *Pessimism* as a control variable does not alter the substantive results for the main explanatory variables and there is no statistically significant relationship between *Pessimism* and conflict. This is consistent with bargaining theory which emphasizes uncertainty in estimates, not the values of estimates themselves, as a cause of war.

Fourth, I again find little evidence consistent with a bureaucratic politics model of communication and information aggregation. If such models are accurate, State Department officials being more optimistic relative to those from other agencies should be associated with a higher probability of using force. Because these officials are expected to be pessimistic, it is only when they go against type that there message is integrated into the final estimate. Department of Defense officials should have the inverse relationship where their pessimistic assessments should reduce the likelihood of using force because it counters expectations. However, neither of these appear in the data. These findings are inconsistent with a cheap talk model of information aggregation and instead support my assumption of preference alignment and credible communication.

[38] All findings are similar without pruning the dictionary.

External Validity

The US, particularly in the data's temporal span, enjoyed a privileged position that is almost unique in the modern state system. Fortuitous geography reduces the importance of defense in US security, making nearly every use of force by the US discretionary and offensive. Extrapolating findings from the US context requires caution on at least two fronts. First, are decision-making processes in other states similarly susceptible to perceptual and overprecision errors? There is reason to believe that such errors are even more pronounced in non-democratic settings. Reiter and Stam (2002, p. 23) theorize that democracies typically make better estimates about relative capabilities than autocracies partially due to a functioning marketplace of ideas.

Second, are diplomats in other states similarly disposed to recognize more uncertainty? I anticipate diplomats have similar dispositions and effects in other countries due to their institutional mandate and expertise about opponent political attributes. This result would attenuate in states where other bureaucracies are equally attendant to the complexity of political attributes and the uncertainty they introduce into estimates of a conflict's expected outcome.

7 Conclusion

This paper demonstrates that incorporating well-established behavioral and psychological factors into rationalist models of war produces novel implications. Once accounting for the possibility of overprecision, certainty is no longer strictly associated with peace. Rather, certainty is problematic when arising from overprecision. I theorize conditions most likely to lead to overprecision in estimates. This stands in contrast to much of the existing literature which presents a list of cognitive biases that impede estimation without placing them in context. The additions to extant models recognize that subjective, and often flawed, estimation is inherent to international security.

Empirically, this paper contributes an original data set. A systematic analysis of declassified security documents sheds light on internal deliberations and the differences across crises, bureau-

cracies, and administrations in a way that is otherwise unavailable. These texts come from the locus of decision-making in US foreign policy and provide a direct measure of the theoretical variables of interest. Statistical analyses indicate that hypotheses integrating rationalist and behavioralist approaches do a better job explaining certainty's observed effect on conflict onset.

A theoretical implication of these results is that asymmetric information is even more problematic than extant research suggests. Incomplete information can generate war through the standard risk-return tradeoff. However, I show it can also generate war due to elites failing to grasp how little information they actually have. An inability to recognize the incompleteness of available information may produce unwarranted certainty and the associated increase in conflict likelihood. This may apply in other political contexts beyond interstate conflict. Bargaining and asymmetric information are fixtures of many political processes. Overprecision in such processes would similarly impede efficiency-improving negotiated outcomes.

Finally, there is a clear policy implication. Estimative errors occur even when the stakes are life and death. This paper specifies when leaders and advisors are most prone to making these errors. Downplaying the importance of understanding an opponent's political attributes by marginalizing those with relevant expertise from the deliberative process increases the risk of these costly misperceptions. Errors of this form account for some of the great failures of US foreign policy in the past half century. Avoiding future costly mistakes may hinge on the ability of leaders and their advisors to recognize the full spectrum of factors that complicate efforts to predict the consequences of using military force.

References

Allison, Graham T. 1971. *Essence of Decision: Explaining the Cuban Missile Crisis*. Boston: Little, Brown.

Alpert, Marc and Howard Raiffa. 1982. A Progress Report on the Training of Probability Assessors. In *Judgment under Uncertainty: Heuristics and Biases*, ed. Daniel Kahneman, Paul Slovic and Amos Tversky. Cambridge: Cambridge University Press.

Arena, Phil and Scott Wolford. 2012. "Arms, Intelligence, and War." *International Studies Quarterly* 56(2):351–365.

Badie, Dina. 2010. "Groupthink, Iraq, and the War on Terror: Explaining US Policy Shift toward Iraq." *Foreign Policy Analysis* 6:277–296.

Bagnoli, Mark and Ted Bergstrom. 1989. "Log-Concave Probability and its Applications." Discussion Paper 89-23, University of Michigan.

Bas, Muhammet A. and Robert J. Schub. forthcoming. "How Uncertainty about War Outcomes Affects War Onset." *Journal of Conflict Resolution* .

Bas, Muhammet and Robert Schub. N.d. "Mutual Optimism as a Cause of Conflict: Secret Alliances and Conflict Onset." Working paper.

Bendor, Jonathan, Ami Glazer and Thomas H. Hammond. 2001. "Theories of Delegation." *Annual Review of Political Science* 4:235–269.

Bendor, Jonathan, Daniel Diermeier, David A. Siegel and Michael M. Ting. 2011. *A Behavioral Theory of Elections*. Princeton, NJ: Princeton University Press.

Bendor, Jonathan and Thomas H. Hammond. 1992. "Rethinking Allison's Models." *American Political Science Review* 86(2):301–322.

Bennett, D. Scott and Allan C. Stam. 2004. *The Behavioral Origins of War*. Ann Arbor, MI: University of Michigan Press.

Bennett, Scott and Allan Stam. 2006. "Predicting the Length of the 2003 US-Iraq War." *Foreign Policy Analysis* 2:101–116.

Betts, Richard K. 1978. "Analysis, War, and Decision: Why Intelligence Failures are Inevitable." *World Politics* 31(1):61–89.

Biddle, Stephen. 2004. *Military Power: Explaining Victory and Defeat in Modern Battle*. Princeton, NJ: Princeton University Press.

Blainey, Geoffrey. 1988. *The Causes of War*. London, Macmillan.

Blechman, Barry and Stephen Kaplan. 1978. *Force without War: US Armed Forces as a Political Instrument*. Washington, DC: Brookings Institution.

Bradley, Ralph A. and Milton E. Terry. 1952. "Rank Analysis of Incomplete Block Designs I: The Method Rank Analysis of Incomplete Block Designs I: The Method of Paired Comparisons." *Biometrika* 39.

Brecher, Michael and Jonathan Wilkenfeld. 1998. *International Crisis Behavior Project, 1918-1994*. Inter-University Consortium for Political and Social Research.

Chiozza, Giacomo and H.E. Goemans. 2011. *Leaders and International Conflict*. Cambridge: Cambridge University Press.

Coelho, Marta P. 2010. "Unrealistic Optimism: Still a Neglected Trait." *Journal of Business Psychology* 25:397–408.

Crawford, Vincent P. and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50(6):1431–1451.

Fearon, James D. 1994. "Political Audiences and the Escalation of International Disputes." *The American Political Science Review* 88(3).

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

Feaver, Peter. 1999. "Civil-Military Relations." *Annual Review of Political Science* 2:211–241.

Fey, Mark and Kristopher W. Ramsay. 2011. "Uncertainty and Incentives in Crisis Bargaining: Game-Free Analysis of International Conflict." *American Journal of Political Science* 55(1):146–169.

Firth, David. 1993. "Bias Reduction of Maximum Likelihood Estimates." *Biometrika* 80(1):27–38.

Fisher, Ronald A. 1935. *The Design of Experiments*. London: Oliver and Boyd.

Fordham, Benjamin. 1998. "The Politics of Threat Perception and the Use of Force: A Political Economy Model of U.S. Uses of Force, 1949-1994." *International Studies Quarterly* 42(3):567–590.

Friedman, Jeff and Richard Zeckhauser. 2012. "Assessing Uncertainty in Intelligence." *Intelligence and National Security* 27(5).

Friedman, Jeffrey A. and Richard Zeckhauser. 2015. "Handling and Mishandling Estimative Probability: Likelihood, Confidence, and the Search for Bin Laden." *Intelligence and National Security* 30(1):77–99.

Fudenberg, Drew and Jean Tirole. 1991. *Game Theory*. Cambridge, MA: MIT Press.

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

Gartzke, Erik and Paul Poast. N.d. "Empirical Implications of Bargaining Theory: An Agenda for Research." Working paper.

Gelman, Andrew. 2006. "The Boxer, the Wrestler, and the Coin Flip: A Paradox of Robust Bayesian Inference and Belief Functions." *The American Statistician* 60(2):146–150.

Gerber, Alan and Donald Green. 1999. "Misperceptions about Perceptual Bias." *Annual Review of Political Science* 2:189–210.

Ghosn, Faten, Glenn Palmer and Stuart Bremer. 2004. "The MID3 Data Set, 1993-2001: Procedures, Coding Rules, and Description." *Conflict Management and Peace Science* 21(2):133–54.

Glaeser, Edward L. and Cass R. Sunstein. 2009. "Extremism and Social Learning." *Journal of Legal Analysis* 1(1):263–324.

Gleditsch, Kristian and Michael D. War. 2001. "Measuring Space: A Minimum Distance Database." *Journal of Peace Research* 38(6):749–768.

Glynn, Adam and Nahomi Ichino. forthcoming. "Using Qualitative Information to Improve Causal Inference." *American Journal of Political Science* .

Grimmer, Justin and Brandon Stewart. 2013. "The Promise and Pitfalls of Automatic Content Analysis Methods for Political Text." *Political Analysis* 3(267-297).

Hafner-Burton, Emilie M., Brad L. LeVeck, David G. Victor and James H. Fowler. 2014. "Decision Maker Preferences for International Legal Cooperation." *International Organization* 68(4):845–876.

Halperin, Morton H. and Priscilla A. Clapp. 2006. *Bureaucratic Politics and Foreign Policy*. Brookings Institution Press.

Hart, Robert A. and David H. Clark. 1999. "Does Size Matter? Exploring the Small Sample Properties of Maximum Likelihood Estimation." Prepared for the Midwest Political Science Association Conference.

Hermann, Margaret G. 1980. "Explaining Foreign Policy Behavior Using the Personal Characteristics of Political Leaders." *International Studies Quarterly* 24(1):7–46.

Horowitz, Michael C. and Allan C. Stam. 2014. "How Prior Military Experience Influences the Future Militarized Behavior of Leaders." *International Organization* 68(3):527–559.

Horowitz, Michael, Rose McDermott and Allan C. Stam. 2005. "Leader Age, Regime Type, and Violent International Relations." *Journal of Conflict Resolution* 49(5):661–685.

Janis, Iriving L. 1982. *Groupthink: Psychological Studies of Policy Decisions and Fiascoes*. Boston: Houghton Mifflin.

Jervis, Robert. 1976. *Perception and Misperception in International Politics*. Princeton, N.J.: Princeton University Press.

Johnson, Dominic. 2004. *Overconfidence and War: The Havoc and Glory of Positive Illusions*. Cambridge, MA: Harvard University Press.

Johnson, Dominic D.P. and Dominic Tierney. 2011. "The Rubicon Theory of War: How the Path to Conflict Reaches the Point of No Return." *International Security* 36(1):7–40.

Jost, John T., Jaime L. Napier, Hulda Thorisdottir, Samuel D. Gosling, Tibor P. Palfai and Brian Ostafin. 2007. "Are Needs to Manage Uncertianty and Threat Associated with Political Conservatism or Ideologicy Extremity." *Personality and Social Psyhcology Bulletin* 33(7):989–1007.

Kahneman, Daniel. 2011. *Thinking Fast and Slow*. Farrar, Strauss and Giroux.

Kahneman, Daniel and Jonathan Renshon. 2007. "Why Hawks Win." *Foreign Policy* 158:34–38.

Kaplow, Jeff and Erik Gartzke. 2013. "Testing the Role of Uncertainty in Conflict." Presented at the 2013 Annual Meeting of the American Political Science Association, Chicago, IL.

Kier, Elizabeth. 1995. "Culture and Military Doctrine: France between the Wars." *International Security* 19(4):65–93.

Krasner, Stephen. 1972. "Are Bureaucracies Important? (Or Allison Wonderland)." *Foreign Policy* 7:159–179.

Krehbiel, Keith. 1991. *Information and Legislative Organization*. Ann Arbor, MI: University of Michigan Press.

Kurizaki, Shuhei. N.d. "Signaling and Perception in International Crises: Two Approaches." Working paper.

Kydd, Andrew. 1997. "Game Theory and the Spiral Model." *World Politics* 49(3):371–400.

Lake, David A. 2010. "Two Cheers for Bargaining Theory: Assessing Rationalist Explanations for the Iraq War." *International Security* 35(3):7–52.

Marsh, Kevin. 2014. "Obama's Surge: A Bureaucratic Politics Analysis of the Decision to Order a Troop Surge in the Afghanistan War." *Foreign Policy Analysis* 10(3):265–288.

Marshall, Monty G. and Keith Jaggers. 2002. *Polity IV Dataset [Computer file]*. College Park, MD: University of Maryland.

McCaskey, Kelly and Carlisle Rainey. N.d. "Logisitic Regression with Small Samples: (Nearly) Unbiased Estimation with Penalized Maximum Likelihood." Working paper.

McManus, Roseanne. forthcoming. "Fighting Words: The Effectiveness of Statements of Resolve in International Conflict." *Journal of Peace Research* .

Mearsheimer, John J. 2001. *The Tragedy of Great Power Politics*. Norton.

Meernik, James. 2004. *The Political Use of Military Force in US Foreign Policy*. Ashgate Press.

Meirowitz, Adam. 2006. "Designing Institutions to Aggregate Preferences and Information." *Quarterly Journal of Political Science* 1:373–392.

Mercer, Jonathan. 2005. "Rationality and Psychology in International Politics." *International Organization* 59(1):77–106.

Milner, Helen and Dustin Tingley. 2015. *Domestic Politics and American Foreign Policy: Sailing the Water's Edge*. Princeton, NJ: Princeton University Press.

Minozzi, William. 2013. "Endogenous Beliefs in Models of Politics." *American Journal of Political Science* 57(3):566–581.

Mitzen, Jennifer and Randall L. Schweller. 2011. "Knowing the Unknown Unknowns: Misplaced Certainty and the Onset of War." *Security Studies* 20:2–35.

Montgomery, Evan Braden. 2006. "Breaking out of the Security Dilemma: Realism, Reassurance, and the Problem of Uncertainty." *International Security* 31(2):151–185.

Moore, Don A. and Paul J. Healy. 2008. "The Trouble with Overconfidence." *Psychological Review* 115(2):502–517.

National Intelligence Estimate. 2002. Iraq's Continuing Programs for Weapons of Mass Destruction.

Oneal, John R., Bruce Russett and Michael L. Berbaum. 2003. "Causes of Peace: Democracy, Interdependence, and International Organizations, 1885-1992." *International Studies Quarterly* 47(3):371–393.

Ortoleva, Pietro and Erik Snowberg. 2015. "Overconfidence in Political Behavior." *American Economic Review* 105(2):504–535.

Potter, Philip. 2007. "Does Experience Matter? American Presidential Experience, Age, and International Conflict." *Journal of Conflict Resolution* 51(3):351–378.

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

Rathbun, Brian C. 2007. "Uncertain about Uncertainty: Understanding the Multiple Meanings of a Crucial Concept in International Relations Theory." *International Studies Quarterly* 51:533–557.

Rathbun, Brian C. 2012. *Trust in International Cooperation International Security Institutions, Domestic Politics and American Multilateralism*. Cambridge: Cambridge University Press.

Rathbun, Brian C. 2015. "Blood, Toil, Tears, and Sweat: Rationality and Romanticism in 1930s British Foreign Policy." Working paper.

Reed, William. 2003. "Information, Power, and War." *American Political Science Review* 97(4):633–41.

Reiter, Dan and Allan C. Stam. 2002. *Democracies at War*. Princeton, NJ: Princeton University Press.

Rosenbaum, Paul R. 2002. *Observational Studies*. New York, NY: Springer-Verlag.

Rovner, Joshua. 2011. *Fixing the Facts: National Security and the Politics of Intelligence*. Ithaca, NY: Cornell University Press.

Rumsfeld, Donald. 2003. "Rumsfeld Foresees Swift Iraq War." Quoted by *BBC News*.

Saunders, Elizabeth N. 2011. *Leaders at War: How Presidents Shape Military Interventions*. Cornell University Press.

Schelling, Thomas C. 1966. *Arms and Influence*. New Haven, CT: Yale University Press.

Schultz, Kenneth A. 1999. "Do Democratic Institutions Constrain or Inform? Constrasting Two Institutional Perspectives on Democracy and War." *International Organization* 53(2):233–266.

Sechser, Todd S. 2011. "Militarized Compellent Threats, 1918-2001." *Conflict Management and Peace Science* 28(4):377–401.

Sekhon, Jasjeet S. 2011. "Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R." *Journal of Statistical Software* 42(7).

Shanker, Thom. 2007. "New Strategy Vindicates Ex-Army Chief Shinseki." *New York Times* 12 January.

Simon, Herbert A. 1955. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics* 69(1):99–118.

Singer, J. David, Stuart Bremer and John Stuckey. 1972. Capability Distribution, Uncertainty, and Major Power War, 1820-1965. In *Peace, War, and Numbers*, ed. Bruce Russett. Sage: Beverly Hills.

Slantchev, Branislav L. and Ahmer Tarar. 2011. "Mutual Optimism as a Rationalist Explanation for War." *American Journal of Political Science* 55(1):135–148.

Smith, Alastair and Allan C. Stam. 2004. "Bargaining and the Nature of War." *Journal of Conflict Resolution* 48(6):783–813.

Smith, Alastair and Allan C. Stam. 2006. "Divergent Beliefs in 'Bargaining and the Nature of War'." *Journal of Conflict Resolution* 50(4):614–618.

Stanovich, Keith E. and Richard F. West. 1998. "Individual Differences in Rational Thought." *Journal of Experimental Psychology* 127(2):161–188.

Teigen, Karl and Magne Jorgensen. 2005. "When 90% confidence intervals are 50% certain: on the credibility of credible intervals." *Applied Cognitive Psychology* 19(4):455–475.

Tetlock, Philip E. 2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton, NJ: Princeton University Press.

Turner, Heather and David Firth. 2012. "Bradley-Terry Models in R: The BradleyTerry2 Package." *Journal of Statistical Software* 48(9):1–21.

Tversky, Amos. 1974. "Assessing Uncertainty." *Journal of the Royal Statistical Society B* 36(2):148–159.

Tversky, Amos and Daniel Kahneman. 1982. Judgment under Uncertainty: Heuristics and Biases. In *Judgment under Uncertainty: Heuristics and Biases*, ed. Daniel Kahneman, Paul Slovic and Amos Tversky. Cambridge: Cambridge University Press.

Van Evera, Stephen. 1999. *Causes of War: Power and the Roots of Conflict*. Ithaca, NY: Cornell University Press.

Waltz, Kenneth N. 1979. *Theory of International Politics*. Addison-Wesley Pub. Co.

Welch, David A. 1992. "The Organizational Process and Bureaucratic Politics Paradigms: Retrospect and Prospect." *International Security* 17(2):112–146.

Yarhi-Milo, Keren. 2013. "In the Eye of the Beholder: How Leaders and Intelligence Organizations Assess Intentions." *International Security* 38(1):7–51.

Appendix 1: Descriptive Statistics and Robustness Tests

Table A1: Summary statistics

	Mean	Std. Dev.	Min.	Max.
Outcome Variable				
Conflict	0.26	0.44	0	1
Explanatory Variables				
Certainty	4.535	1.117	1	6.468
State Dept Involvement	0.51	0.50	0	1
Control Variables				
Relative Capabilities	0.90	0.16	0.45	0.99
Regime Type	-5.77	3.84	-10	10
Proximity	9,719	4,114	1,813	16,337
Non-State Enemy	0.21	0.41	0	1
Republican	0.51	0.50	0	1
Combat	0.72	0.45	0	1
Age	57.2	7.6	44	70
Experience	1.9	1.8	0	7
Pessimism	0.144	0.027	0.097	0.217
N	61	61	61	61

Table A2: Linear Probability Models

	(1)	(2)	(3)	(4)
Certainty	-0.04	-0.06	0.17**	0.09
	(0.05)	(0.05)	(0.08)	(0.09)
State Dept Involvement		-0.23*	1.30***	0.81
		(0.13)	(0.46)	(0.52)
Certainty*State Dept	—	—	-0.32***	-0.23**
			(0.10)	(0.11)
Combat		-0.11		-0.12
		(0.13)		(0.12)
Relative Capabilities		0.95**		0.72*
		(0.38)		(0.38)
Regime Type		0.03**		0.03**
		(0.02)		(0.02)
Proximity		0.00		0.00
		(0.00)		(0.00)
Republican		-0.05		-0.07
		(0.17)		(0.16)
Age		0.00		0.00
		(0.01)		(0.01)
Non-State Enemy		-0.14		-0.08
		(0.16)		(0.16)
Constant	0.44*	-0.24	-0.41	-0.47
	(0.24)	(0.69)	(0.36)	(0.68)
N	61	61	61	61

$^{*}p < 0.1$ $^{**}p < 0.05$ $^{***}p < 0.01$

Notes: Ordinary least squares regression with opportunities to use force as unit of analysis.

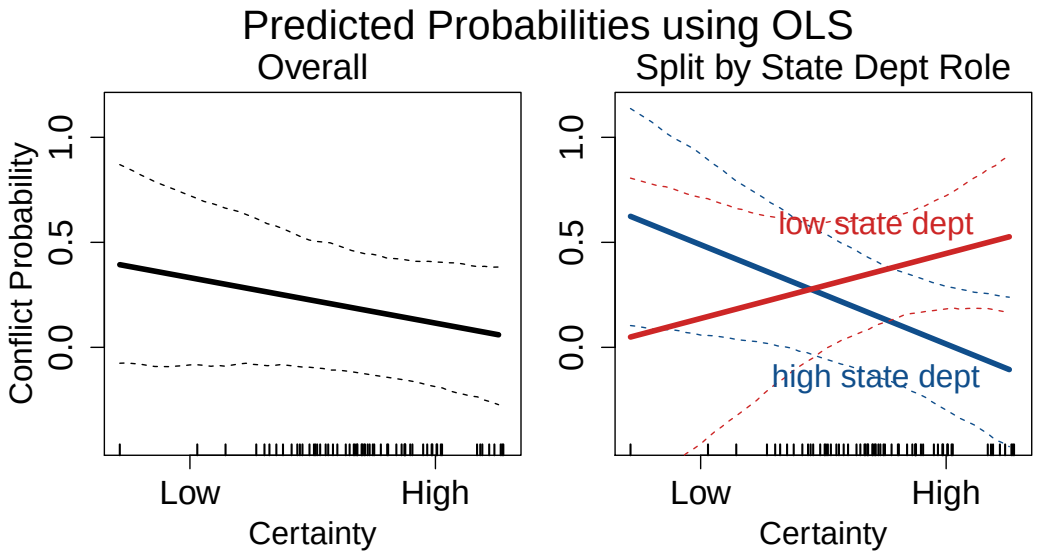

Figure A1: Predicted probabilities of conflict across the range of certainty values with 95% confidence intervals. Left panel based on Model 2 from the table of linear probability models. Right panel based on Model 4 from the table of linear probability models. Rug plot shows data distribution.

Table A3: Penalized Maximum Likelihood

	(1)	(2)	(3)	(4)
Certainty	-0.19	-0.45	0.74*	0.23
	(0.25)	(0.30)	(0.44)	(0.51)
State Dept Involvement		-1.43*	6.60**	3.56
		(0.77)	(2.84)	(3.08)
Certainty*State Dept	—	—	-1.66***	-1.13*
			(0.63)	(0.69)
Combat		-0.50		-0.61
		(0.84)		(0.90)
Relative Capabilities		8.67*		6.86
		(4.98)		(4.95)
Regime Type		0.18*		0.19*
		(0.09)		(0.10)
Proximity		0.00		0.00
		(0.00)		(0.00)
Republican		-0.14		-0.18
		(1.00)		(1.06)
Age		-0.01		-0.01
		(0.07)		(0.07)
Non-State Enemy		-0.63		-0.27
		(1.07)		(1.09)
Constant	-0.16	-5.25	-4.05*	-6.12
	(1.14)	(5.49)	(2.12)	(6.04)
N	61	61	61	61

$^{*}p < 0.1$ $^{**}p < 0.05$ $^{***}p < 0.01$

Notes: Penalized maximum likelihood (Firth 1993) with opportunities to use force as unit of analysis.

Table A4: Alternative Control Variables

	(1)	(2)	(3)	(4)	(5)
Certainty	0.33	0.28	0.34	0.32	0.39
	(0.60)	(0.62)	(0.60)	(0.61)	(0.62)
State Dept Involvement	5.71	5.89	5.76	5.13	6.01
	(3.84)	(3.95)	(3.85)	(3.93)	(3.91)
Certainty*State Dept	-1.74**	-1.75**	-1.74**	-1.62*	-1.73**
	(0.88)	(0.88)	(0.87)	(0.88)	(0.88)
Combat	-0.75	-0.80	-0.75	-1.75	-0.36
	(1.06)	(1.11)	(1.08)	(2.00)	(1.26)
Relative Capabilities	19.13	21.82	20.42	19.35	16.62
	(19.54)	(24.17)	(22.10)	(18.41)	(17.77)
Regime Type	0.26**	0.26**	0.26*	0.31*	0.27**
	(0.13)	(0.13)	(0.13)	(0.16)	(0.13)
Proximity	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Republican	-0.24	0.00	-0.09	-0.69	0.04
	(1.31)	(1.29)	(1.31)	(1.54)	(1.31)
Age	0.00	-0.02	-0.02	0.00	-0.01
	(0.12)	(0.09)	(0.08)	(0.09)	(0.08)
Non-State Enemy	-0.71	-0.68	-0.75	-0.94	-0.86
	(1.41)	(1.38)	(1.39)	(1.43)	(1.38)
Experience	-0.06				
	(0.32)				
Pessimism		51.39			
		(56.83)			
Speakers			0.02		
			(0.08)		
Time				-0.08	
				(0.13)	
Recent War					-0.05
					(0.09)
Constant	-19.11	-22.18	-19.69	134.85	-17.07
	(20.07)	(24.81)	(22.69)	(251.57)	(18.15)
N	61	61	61	61	61

$^{*}p < 0.1$ $^{**}p < 0.05$ $^{***}p < 0.01$

Notes: Logistic regression with opportunities to use force as unit of analysis. Model 1 controls for the president's time in office, Model 2 for the pruned pessimism dictionary score, Model 3 for the number of speakers, Model 4 for time, and Model 5 for time since the last major US war.

Table A5: Alternative Standard Error Corrections

	Non-Bootstrapped			Bootstrapped		
	No Cluster	Admin	Enemy	No Cluster	Admin	Enemy
Certainty	0.09	0.09	0.09*	0.09	0.09	0.09
	(0.09)	(0.08)	(0.05)	(0.08)	(0.08)	(0.07)
State Dept Involvement	0.81	0.81**	0.81**	0.81	0.81**	0.81*
	(0.52)	(0.29)	(0.34)	(0.56)	(0.37)	(0.43)
Certainty*State Dept	-0.23**	-0.23**	-0.23***	-0.23*	-0.23***	-0.23**
	(0.11)	(0.07)	(0.08)	(0.12)	(0.08)	(0.10)
Combat	-0.12	-0.12*	-0.12	-0.12	-0.12	-0.12
	(0.12)	(0.05)	(0.12)	(0.12)	(0.20)	(0.13)
Relative Capabilities	0.72*	0.72***	0.72*	0.72**	0.72**	0.72
	(0.38)	(0.14)	(0.41)	(0.31)	(0.31)	(2.27)
Regime Type	0.03**	0.03*	0.03**	0.03*	0.03**	0.03
	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Proximity	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Republican	-0.07	-0.07	-0.07	-0.07	-0.07	-0.07
	(0.16)	(0.06)	(0.13)	(0.18)	(0.28)	(0.17)
Age	0.00	0.00	0.00	0.00	0.00	0.00
	(0.01)	(0.00)	(0.01)	(0.01)	(0.02)	(0.01)
Non-State Enemy	-0.08	-0.08	-0.08	-0.08	-0.08	-0.08
	(0.16)	(0.05)	(0.14)	(0.16)	(0.10)	(0.16)
Constant	-0.47	-0.47	-0.47	-0.47	-0.47	-0.47
	(0.68)	(0.35)	(0.48)	(0.63)	(1.13)	(2.11)
N	61	61	61	61	61	61

$^{*}p < 0.1$ $^{**}p < 0.05$ $^{***}p < 0.01$

Notes: Ordinary least squares regression with opportunities to use force as unit of analysis and alternative standard error corrections. The first three models do not use bootstrapped standard errors while the second three models do. The models vary in whether standard errors are clustered, and if so, on what variable as indicated in model title. They are not clustered, clustered on the presidential administration, or clustered on the opponent.

Table A6: Excluding Repeated Observations from Vietnam War

	(1)	(2)	(3)	(4)
Certainty	1.14*	1.19	0.66	1.55
	(0.66)	(0.98)	(1.04)	(1.25)
State Dept Involvement	9.35**	9.48*	13.29*	16.92**
	(3.84)	(5.42)	(7.99)	(8.13)
Certainty*State Dept	-2.21***	-2.68**	-3.55*	-4.29**
	(0.84)	(1.25)	(1.84)	(1.89)
Combat		-0.03	1.10	1.12
		(1.30)	(1.51)	(1.54)
Relative Capabilities		16.99	16.22	15.59
		(20.12)	(18.34)	(14.14)
Regime Type		0.28*	0.32	0.25
		(0.15)	(0.20)	(0.19)
Proximity		-0.00	0.00	-0.00
		(0.00)	(0.00)	(0.00)
Republican		-1.03	-0.02	0.04
		(1.59)	(1.96)	(1.87)
Age		0.02	-0.09	-0.06
		(0.10)	(0.13)	(0.13)
Non-State Enemy		-0.26	-1.17	-0.62
		(1.43)	(1.58)	(1.62)
Constant	-6.38**	-21.13	-12.58	-18.00
	(3.18)	(21.50)	(20.56)	(18.76)
N	53	53	38	43

$^{*}p < 0.1$ $^{**}p < 0.05$ $^{***}p < 0.01$

Notes: Logistic regression with opportunities to use force as unit of analysis. All models exclude observations from Vietnam War except for the first observation in each administration. Excluding all Vietnam War observations produces similar results. Models 1 and 2 include the full sample of observations. Model 3 includes only those observations in ICB or meeting Blechman and Kaplan's levels one and two uses of force. Model 4 excludes observations with fewer than 500 words of text.

Appendix 2: Sampling Criteria

International Crisis Behavior Data

Crisis defined as a "situation deriving from change in a state's internal or external environment which gives rise to decision makers' perceptions of *threat to basic values, finite time for response*, and the *likelihood of involvement in military hostilities* [emphasis in original text]." From Brecher and Wilkenfield (1982) [p. 383].

Blechman and Kaplan Level of Force

- *Level 1:* Use of strategic nuclear unit plus at least one "major" force component
- *Level 2:* Two or three "major" force components used, but not strategic nuclear units

Where a "major" unit is (1) two or more aircraft carrier task groups, (2) more than one army battalion, or (3) one or more combat wings.

Appendix 3: Dictionary Validation

I approach dictionary method validation in three ways. The first examines speakers or texts that offer a clear relative ranking of their respective uncertainty levels. Beyond the example included in the main text, I repeat the exercise with two texts from during the Pleiku crisis in 1965. Thomas Hughes, Director of the State Department's Bureau of Intelligence and Research, was dismayed by an earlier memo from McGeorge Bundy. In Bundy's memo concerning probable reactions to escalated US bombing in Vietnam, he understated the risk of bombing actions by ignoring probable Chinese reactions. Chinese intervention or support for the North Vietnamese would amount to greater enemy military capabilities and thus is an added source of uncertainty. Hughes writes:

> Incomprehensibly to me, the White House memorandum discusses the risks of sustained US air strikes against North Vietnam without examining Chinese Communist responses. However, the two intelligence community products estimate Chinese Communist air intervention to be quite likely at some stage in this very process.[39]

Again, the measure appropriately identifies Hughes' memo to be more uncertain than Bundy's memo (uncertainty scores of 7.7% and 5.3% respectively).

[39]Memorandum from the Director of the Bureau of Intelligence and Research (Hughes) to Acting Secretary of State Ball, February 8, 1965, *FRUS*, Volume II, Vietnam, 1964-1968, Document 90. Note that this document is not addressed to the president and is therefore only used for validation purposes.

Another validity check compares uncertainty scores of those with historically well-known assessments. For instance, George Ball is widely thought to have been uncertain about what outcomes the US could secure through force in Vietnam. Walt Rostow, in contrast, was an adamant believer that North Vietnamese resolve was limited and the continued application of force would prove effective. In accordance with expectations, the uncertainty score for Ball (5.4%) is higher than that for Rostow (4.6%).

For the second and third validity checks, a research assistant (RA) and I hand coded a subset of documents. For ease of comparability, I use memos in the corpus that are between 50 and 100 words, which provides a subset of 44 documents. Memos are preferable to transcripts for hand coding because transcript speaker texts are removed from the transcript context and consequently substantively harder to interpret as standalone documents. The 50 to 100 word range insures all speaker texts within the subset are relatively comparable and of a length where coders can comprehend the content. The Bradley-Terry scoring approach is based on 381 pairwise comparisons.

The continuous uncertainty score provided by the dictionary method complicates validity assessments as there are not discrete categories to compare to the hand coded categorization scheme. To address this, I dichotomize the dictionary uncertainty scores as either low or not low, where the former category contains texts with uncertainty scores more than one standard deviation below the mean value. Eight of the 44 texts are coded as low uncertainty by this criteria. Using the hand categorization by an RA and myself as a benchmark, the dictionary method accurately classifies 77% and 68% of texts, respectively. Accuracy is defined as texts correctly categorized over total texts. F-scores provide an alternative assessment tool that incorporates both categorization precision and recall. Using the RA's coding as a baseline produces scores of $F_{NL} = .85$ and $F_L = .5$, where the subscripts indicate "not low" uncertainty and "low" uncertainty respectively. Using my own coding produces $F_{NL} = .78$ and $F_L = .42$.

Appendix 4: Randomization Inference

This section describes the calculation for the test statistic used in the randomization inference robustness tests. First, I "residualize" the data using an OLS specification that excludes the main explanatory variables (certainty and State Department involvement) and their interaction. Second, I regress the residuals on the uncertainty measure for each observation with the sample split based on the level of State Department involvement. Third, the test statistic is equal to the difference in the coefficients from these two linear regressions. This test statistic is then compared to those calculated from 5,000 randomizations of the explanatory variables with are randomly assigned

without replacement in each iteration.

Appendix 5: Speakers with Most Text by Administration

Table A7: Influential Speakers by Administration

Eisenhower	Kennedy	Johnson	Nixon	Ford	Carter
Dulles (John)	McNamara	Bundy	Kissinger	Kissinger	Brzezinski
Radford	Taylor	McNamara	Laird	Colby	Turner
Dulles (Allen)	Rusk	Ball	Newsom	Schlesinger	Vance
Cutler	Komer	Rusk	Rogers	Jones	Muskie
Smith (Walter)	Bundy	Rostow	Moorer	Wyand	Brown
Stassen	McCone	McCone	Helms	Rockefeller	Sullivan
Morgan	Schlesinger	Wheeler	Haig	Rumsfeld	Aaron

Notes: Speakers with the most text by administration.

Appendix 6: Uniform Distribution

This section sketches the comparative statics of interest using a uniform distribution (results are substantively similar for beta distributions with non-decreasing hazard rates). Let c_B be drawn from a cumulative distribution $H(z)$. As established by Fearon (1995), in the substantively interesting range of parameter values, A maximizes $u_A(x)$ with $x^* \in [p, 1]$ that solves $\frac{h(x-p)}{1-H(x-p)} = \frac{1}{x-p+c_A}$. I consider the case with $c_B \sim U(\mu - \Delta, \mu + \Delta)$, where μ is the mean of the distribution and Δ represents the degree of uncertainty.

Begin with the case of warranted certainty. Substituting a uniform distribution into the general condition, solve for x^*. This produces $[\frac{1}{2\Delta}](p - c_A - x) + 1 - [\frac{x-p-\mu+\Delta}{2\Delta}]$=0, which reduces to $x^* = \frac{\Delta+2p+\mu-c_A}{2}$. There are two cases. In case (i), $x^* \geq p + \mu - \Delta$, A demands $\frac{\Delta+2p+\mu-c_A}{2}$ which is increasing in Δ. Note that due to the finite bounds of the distribution, A demands at least $p + min(c_B)$ or $p + \mu - \Delta$. In case (ii), $x^* < p + \mu - \Delta$ in which case A demands $p + \mu - \Delta$, which is decreasing in Δ. Thus the relationship is non-monotonic.

Uncertainty's relationship to the probability of war is the main concern. In the uniform case, Pr(War)$= H(x^*-p)$. Substituting in, Pr(War)=$\frac{x^*-p-[\mu-\Delta]}{2\Delta}$. In case (i) this reduces to Pr(War)=$\frac{3\Delta-c_A-\mu}{4\Delta}$ and thus the probability of war is increasing in uncertainty (Δ). In case (ii) varying Δ has no effect because the probability is fixed at zero.

Now consider the case with unwarranted certainty. Let θ represent the unwarranted percentage uncertainty reduction and solve for x^*. Repeating the above analysis produces $[\frac{1}{2(1-\theta)\Delta}](p - c_A - x) + 1 - [\frac{x-p-\mu+(1-\theta)\Delta}{2(1-\theta)\Delta}]$=0, which reduces to $x^* = \frac{\Delta-\theta\Delta+2p+\mu-c_A}{2}$. There are two cases. In case (I) $x^* \geq p + \mu - (1-\theta)\Delta$, A demands $\frac{\Delta-\theta\Delta+2p+\mu-c_A}{2}$ which is decreasing in θ. Again, with finite bounds on the distribution A demands at least $p + min(c_B)$ as $min(c_B)$ is perceived. In case (II) $x^* < p + \mu - (1-\theta)\Delta$, A demands $p + \mu - (1-\theta)\Delta$, which is increasing in θ. The relationship is non-monotonic.

Solving for the probability of war is akin to the analysis above. A demands x^*, which is determined based on the misperceived distribution of c_B while c_B is drawn from its actual full range; Pr(War)= $H(x^* - p)$. In case (I) this reduces to Pr(War)= $\frac{3\Delta-c_A-\mu-\theta\Delta}{4\Delta}$ which is decreasing in unwarranted certainty (θ). In case (II), Pr(War)= $\frac{p+\mu-(1-\theta)\Delta}{2\Delta}$ which is increasing in θ. Certainty, when unwarranted, can increase the probability of war.