

# Leaders, States, and Reputations

Cathy X. Wu<sup>1</sup>   
and Scott Wolford<sup>2</sup>

Journal of Conflict Resolution  
2018, Vol. 62(10) 2087-2117  
© The Author(s) 2018  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/0022002718786001  
journals.sagepub.com/home/jcr



## Abstract

Reputational incentives are ubiquitous explanations for war, yet consistent evidence of their effects is elusive for two reasons. First, most work searches for the payment of reputational costs, yet strategic censoring systematically biases observational data against revealing them. Second, the locus of reputation is often ambiguous, yet the choice of leader or state as unit of observation has inferential consequences. Our research design (a) focuses on observable implications of reputational theories in appropriate samples and (b) considers two competing sources of reputational incentives: changes in national leaders and in political institutions. Consistent with our expectations, leadership turnover and regime change are each associated with initially high probabilities that militarized disputes escalate to the use of force before declining over time in the presence of a reasonable expectation of future disputes. Reputations are in evidence, but analysts must look for them in the right place.

## Keywords

militarized disputes, rivalry, political leadership, bargaining

Leaders often invoke the maintenance of their reputation, credibility, or prestige when explaining decisions to engage in armed conflict, citing the possibility of challenges or even war tomorrow if toughness, commitment, or resolve is not shown

---

<sup>1</sup>Department of Political Science and Geography, Old Dominion University, Norfolk, VA, USA

<sup>2</sup>Department of Government, University of Texas, Austin, TX, USA

## Corresponding Author:

Cathy X. Wu, Department of Political Science and Geography, Old Dominion University, 7000 Batten Arts and Letters, Norfolk, VA 23529, USA.

Email: [xwu@odu.edu](mailto:xwu@odu.edu)

today. But if reputation is “one of the few things worth fighting over” (Schelling 1966, 124), the absence of a consistent body of evidence that decision makers cultivate and attribute reputations is puzzling (see, *inter alia*, Dafoe, Renshon, and Huth 2014; Danilovic 2002; Huth 1997, 1999; Huth and Russett 1993; Mercer 1996; Press 2005; Weisiger and Yarhi-Milo 2015).<sup>1</sup> If leaders care deeply about their own and others’ reputations, why does there appear to be such a tenuous relationship between past behavior and present outcomes in the empirical record? The answer is simple: most research on the question looks in the wrong place.

First, most empirical work adopts a research design that looks for links between past and present behavior and beliefs, for evidence that actors pay reputation costs for backing down or showing weakness (e.g., Mercer 1996; Press 2005; Weisiger and Yarhi-Milo 2015). But if leaders wish to avoid paying them, the largest such costs are unlikely to be paid on the equilibrium path—that is, in the observational data that constitute the empirical record. Reputations can shape patterns of war and peace, yet produce no obvious links between past and present outcomes in observational data. Recent theoretical models have developed strategic links between private information, reputation-building, and war (Fearon 1995; Nalebuff 1991), leading to the generation of additional empirical implications (see, e.g., Wolford 2007), but attempts to translate these insights into more appropriate research designs are few (but see Clare and Danilovic 2010, 2012; Crescenzi 2007; Rider 2013; Spaniel and Smith 2015). In contrast, we derive hypotheses over outcomes that occur not *off* the equilibrium path but *on* it, over relationships present in observational data: temporal cycles of dispute escalation driven by the introduction of new private information between states that expect to bargain with one another both today and into the future.

Second, research designs often confuse two potential loci of reputations (cf. Huth 1997, 78–80), political regimes and national leaders, yet choosing units of analysis has implications for what constitutes evidence of behavior driven by reputations. Press (2005), for example, argues that aggressive Soviet demands in the Berlin Crisis of 1961 were not driven by prior American behavior—that is, Eisenhower’s 1958 demonstrated commitment to West Berlin—as they should if “past actions” shaped the credibility of present threats.<sup>2</sup> Yet both Soviet behavior at the Vienna Summit and the American response *are* consistent with a reputation account, if successive leaders in the same state can differ in their preferences over the resort to war (Wolford 2007, 772). If past American behavior mattered little for Soviet decision-making, then it may have been due to the accession of a new president in Kennedy with uncertain preferences and reputation incentives to misrepresent them. The heightened risk of war in March 1961 is inconsistent with a state-centric view but consistent with a leader-centric one. This does not imply that reputations for resolve inhere in either political regimes or leaders alone, only that multiple loci of reputation might both exist. Nonetheless, most research designs elide this distinction.

Reputations emerge from private information over some underlying disposition and incentives to shape others' beliefs about it, but reputations are worth fighting over only when they travel across crises and issues. One such piece of private information is a state's willingness to use force, which depends on a leader's subjective estimate of the costs of fighting (her resolve), modified by any preferences induced by the political institutions that constrain her choices (cf. Bueno de Mesquita et al. 2003).<sup>3</sup> Observable background experiences shape a leader's resolve in part (Horowitz et al. 2018), but unobservable attributes and beliefs, such as her ability to mobilize domestic support, her relationships with advisors (Saunders 2018), and her view of institutional constraints (Clark and Nordstrom 2005), create uncertainty over her willingness to use force. And since irresolute leaders have strategic incentives to misrepresent their preferences (Fearon 1995, 395-401), changes in either national leadership *or* political institutions can introduce new information problems into a country's foreign relations. Leader and regime changes thus create "informational traps" in which both sides have strategic incentives—one to convey information and the other to gain it—to take steps that increase the chances of war. Those chances then fall with the passage of time and the revelation of information, whether a leader reveals an abundance *or* a lack of resolve.

The changes associated with these informational traps—the accession of new national leaders and significant changes in political institutions—can occur simultaneously. To address this inferential problem, we estimate empirical models designed to parse the independent effects of reputations that inhere in leaders and political regimes, using indicators of the time since the introduction of each type of private information to explain changes in the probability of dispute escalation. Reputation incentives should only emerge in states that expect future contentious interactions, so we use several samples, including strategic rivals (Thompson 2001; Thompson and Dreyer 2012), contiguous states (Stinnett et al. 2002), and politically relevant dyads involving both neighbors and major powers (Lemke and Reed 2001) from 1855 to 2007, where reputation incentives are at their strongest. Consistent with the expectations derived from our theoretical model, the probability that disputes escalate to the use of force is highest when leaders *or* political regimes are new—that is, when informational traps are difficult to escape. These patterns are clearest in rival, contiguous, and politically relevant dyads, while they disappear in larger samples that include states with little or no reputation stakes in their interactions.

Evidence of the cultivation and attribution of reputations *can* be found, but analysts must look in the right place. The empirical record is systematically unlikely to produce the most commonly sought evidence, so analysts must match observational data to hypotheses about behavior that occurs on the equilibrium path—that is, to observable implications. We find that reputations are in evidence for both leaders and political regimes, which suggests that ongoing debates over the level of analysis problem in international relations may be counterproductive (see Braumoeller 2012; Chiozza and Goemans 2004a; Horowitz, Stam, and Ellis 2015; McGillivray and Smith 2008; Singer 1961). Both national leaders and political regimes appear

capable of developing reputations, which implies that a strict focus on one level of analysis to the exclusion of the other can mask critical relationships between leader tenure, regime survival, and the outbreak of war.

## **Reputation and Crisis Bargaining**

Before discussing the literature, we clarify the observable implications of reputations for crisis bargaining. Our game-theoretic model (analyzed in the Online Appendix) shows that the clearest link between past and present crisis behavior emerges not in the payment of reputation costs, not in the onset of new crises, but in the escalation of crises to war in observational data. The reputation costs for which many studies search *do* influence observable behavior, but two factors ensure that these costs are largely unobservable in the historical record. First, actors try to avoid the highest reputation costs, keeping the most easily measured consequences of showing weakness in a crisis off the equilibrium path (cf. Schultz 2001). Second, such costs players are willing to pay should manifest not in an elevated probability of war during subsequent crises but in the bargains on offer, where the information gleaned from a revelation of dovishness has the same impact as a revelation of hawkishness: a reduction in the probability that crises escalate to war (cf. Wolford 2007, 779-80). In contrast to the payment of reputation costs, these temporal cycles of dispute escalation are a directly observable implication of the cultivation and attribution of reputations, and our theoretical model allows us to derive expectations over where and how to find them.

We define player *B*'s reputation as the subjective belief held by player *A* about some dispositional but unobservable characteristic that *B* carries across different strategic situations (Kreps and Wilson 1982; Milgrom and Roberts 1982; Nalebuff 1991; Sartori 2005).<sup>4</sup> Reputations may form around the value *B* places on honoring commitments (George and Smoke 1974; Schelling 1966), the capabilities or effectiveness of *B*'s military (Huth 1997, 76-77), or *B*'s subjective assessment of the costs of using military force to achieve his ends—that is, his resolve.<sup>5</sup> The latter shapes *B*'s subjective assessment of the desirability of using force in general (cf. Nalebuff 1991; Sartori 2005); when his costs are high, *B* is irresolute (or dovish), and when his costs are low, *B* is resolute (or hawkish). Finally, *B*'s privately known preferences are unobservable, forcing *A* to use her prior beliefs and *B*'s observable actions to form her own estimate of *B*'s resolve. This estimate, held by *A* and based in part on *B*'s actions, constitutes *B*'s reputation.

Reputations play an important role in models of crisis bargaining under asymmetric information because they inform the prior beliefs with which players enter the game. Beliefs over which bargains each side will accept shape crisis outcomes, giving *B* an incentive to misrepresent his resolve today, if he can do so cheaply, to influence the outcomes of future crises. In other words, he'd lie if he could. This undermines *B*'s ability to convey information about his resolve even if he is truly willing to fight (Fearon 1995, 390-401); incentives to misrepresent create a

corresponding incentive to disbelieve. Conveying resolve credibly requires that a signal be sufficiently costly—financially, militarily, emotionally, or politically—to deter irresolute types from mimicking the signal in attempts to appear resolute (Fearon 1997; Jervis 1970; Spence 1973). And in crisis bargaining, fighting or risking war today can signal resolve where simple verbal statements can't (George and Smoke 1974; Schelling 1966). The costs of fighting, from death and destruction to the disfigurement of the economy and body politic, can signal resolve when players bear them, and avoiding those costs can signal irresolution, which creates a set of dangerous strategic incentives.

In our theoretical model, two players, *A* (she) and *B* (he), confront one another in two sequential disputes that end in war if the players fail to strike a bargain. Like Wolford's (2007) model, *A* begins the game uncertain over *B*'s resolve, modeled as his private costs of war. Where Wolford's model assumes the occurrence of two sequential crises, we generalize it in two variants, giving *A* and *B*, respectively, the choice to initiate a dispute in the first period or to pass and wait until the second period, where a second dispute occurs with probability  $\delta$ —a common representation of the shadow of the future (Powell 1999). This allows us to track, first, how the probability of war changes from one dispute to the next as a function of information revealed (i.e., reputations attributed) and, second, how *A*'s beliefs about *B*'s resolve affect both players' willingness to initiate disputes. Reputation scholarship rarely differentiates the onset and escalation of disputes, but we show that the distinction turns out to be critical. Our model generates no consistent bivariate relationship between the passage of time and the occurrence of disputes, but it does predict one between the passage of time and the escalation of disputes to the use of force.

In each crisis, *A*'s uncertainty over *B*'s resolve creates a risk-return trade-off (see Powell 1999, chap. 3): the more she demands, the better off she is if *B* accepts, but larger demands also mean that *B* is more likely to find the deal unacceptable and choose war instead. To the extent that *A* believes *B* to be resolute, *A* offers better terms than she would if she believed *B* to be irresolute, and this encourages *B* to cultivate a reputation for resolve by rejecting *A*'s offer in the first crisis. At the same time, *A* wishes to tailor her strategy to receive the best possible bargains over time at the lowest cost and doing so is easier the more she knows about *B*'s resolve. With cheap talk uninformative, *A* makes aggressive demands that force *B* to reveal his resolve or lack thereof. And if *B* is relatively resolute, his best response is to engage in costly or risky behavior of his own—escalating to the use of force, for example—in an effort to convince *A* that he is in fact resolute. When relatively dovish, on the other hand, *B* accepts and earns a correspondingly dovish reputation. *A* thus chooses her demands to ensure that acceptance and rejection both reveal a useful amount of information.<sup>6</sup> Each player has strategic incentives to take costly, risky steps to reveal or conceal private information, increasing the chances of dispute escalation whose costs he or she is bound to regret, creating what we term an "informational trap." Yet even if *B* reveals dovishness, guaranteeing less favorable bargains in the future, the probability of disputes escalating to violence declines for the second crisis, just as it

would if he conveyed hawkishness; useful information is revealed to *A* in either case. And if *A* does not take the chance to learn about *B*'s resolve, she foregoes a better deal in a second crisis with a lower probability of war. This relationship exists when the probability of a future crisis ( $\delta$ ) is sufficiently high, such that the costs of developing a reputation for resolve are worth paying; if *A* and *B* are unlikely to interact in the future, there are no additional gains to be had from risking war today.

When new private information is introduced into this strategic system, each player is trapped, unable to resist taking steps that raise the probability of war, steps that she or he would not countenance if the shadow of the future did not make a reputation for resolve so valuable.<sup>7</sup> Disputes are thus most likely to escalate immediately following a change in *B*'s private information—for us, a change in leadership or political institutions—but the chances of escalation then decline over time as *B* reveals information about his resolve, allowing *A* to tailor bargains that minimize the risks of escalation and ensure a favorable string of bargains at minimal cost. This is our model's primary observable implication, characterized formally by Proposition 3 in the Online Appendix: when future disputes are sufficiently likely, the probability of dispute escalation peaks when *B*'s private information is new, then declines over time, as early interactions reveal information about *B*'s resolve.

Our theory also indicates the most relevant sample for empirical analysis, in that it relies on actors that bargain in the shadow of expected future crises and disputes. Resolve must have later crises to which it can travel if information gleaned or revealed today is not to be a wasting asset, so players hesitate to demonstrate resolve for future crises that will never happen, and the difference in the probability of war between the first and second periods will be smaller.<sup>8</sup> An adequate test of the theory's implications, then, tracks players over time and identifies those sufficiently concerned about future crises to act on reputation incentives in the present. Our research design reflects this key condition by sampling on strategic rivals, that is, states that both view one another as long-term threats and experience multiple crises and disputes (cf. Clare and Danilovic 2010; Thompson 2001), contiguous dyads (Stinnett et al. 2002), and politically relevant dyads (Lemke and Reed 2001), which minimizes the risk of aggregation bias in larger samples that don't fit the theory's assumptions.

The informational value of the first crisis leads to a further implication: a demonstrated lack of resolve in early disputes should *not* be associated in observational data (i.e., in equilibrium) with a subsequently heightened probability of escalation. *A* can calibrate her demands using what she learns from an initial crisis, so the probability of escalation falls over time even if *B* develops a reputation for dovishness: he simply receives less favorable bargains in the future. Some reputation costs are paid on the equilibrium path in the form of less favorable bargains when *B* fails to convey resolve in the first crisis, but they are a small, biased sample of the true range of possible reputation costs (cf. Schultz 2001). If the costs of revealing dovishness are too large, *B* opts instead to mimic resolute types by fighting or risking a war. Thus, any link between past actions and present crisis outcomes should be off the

equilibrium path, systematically minimized in or censored out of observational data. The payment of reputation costs is a causally important element of reputation theories, but it is not an *observable* element.

This leads us to expect no consistent bivariate relationship between the passage of time and the initiation or onset of disputes. The probability that crises escalate does not depend on *what* information is revealed in early interactions, only on the fact that information *is* revealed, but a player's decision to initiate a crisis depends directly on the content of that information. When *A* has the opportunity, initiating a crisis is more attractive when *B*'s reputation is dovish and less so when his reputation is hawkish, but reputations can move in either direction, hawkish or dovish, over time. Likewise, the probability with which *B* himself initiates a crisis, revealing information even before the crisis unfolds, depends on the interaction of his valuation of the status quo and his current reputation. As the latter changes, the probability of crisis onset can rise or fall over time, but the probability that these disputes escalate to war continues to fall from one crisis to the next. Reputation-building can generate *any* relationship—positive, negative, or even null—between the passage of time and the onset of crises, even if analysts have a way of solving problems of strategic censoring. Therefore, crisis onset is a less useful outcome variable than crisis escalation when it comes to looking for evidence of reputations: past and present can be linked in any direction when it comes to the chances of dispute onset, but the effect of time on escalation is clear and unconditional.

If theories of reputation rule out the possibility of observing the largest reputation costs in observational data, then the search for the consequences of “past actions” (e.g., Press 2005) is constrained to be uninformative. Reputation costs may go unpaid because reputations do not matter *or* because such costs are either (a) so large that leaders work very hard to avoid paying them or (b) associated with the same observable outcome as demonstrating resolve: a reduced probability that disputes escalate to war. But by deriving implications over the emergence of informational traps, the passage of time, and the probability of escalation, we *can* arbitrate between claims that reputation incentives do or do not affect processes of crisis bargaining and escalation. If previous interactions provide no important information for current disputes, if past and present are not linked by a process of reputation cultivation and attribution, then elapsed time since changes in actor *B*'s unobservable preferences should not be related to the probability that disputes escalate. If, on the other hand, the probability of escalation peaks after the introduction of new private information and declines over time, we can be more confident that the cultivation and attribution of reputations shape observed patterns of war and peace.

## Looking for Reputations

Reputations have been central to the study of interstate conflict since the beginning of the Cold War (George and Smoke 1974; Huth and Russett 1993; Mearsheimer 1983; Russett 1963; Schelling 1966), as the United States searched for ways to make

credible extended deterrent commitments made in the wake of the Second World War it had studiously avoided making after the First. Yet, despite a long theoretical tradition, consistent evidence of reputations in crisis bargaining remains lacking for two reasons. First, scholarship on the question often chooses problematic research designs, searching for the payment of costs that should be censored in observational data. Second, the same work—like our stylized theoretical model—tends to conflate two distinct loci of reputation: states and their leaders.<sup>9</sup> But both new leadership and new institutional arrangements can introduce consequential private information into crisis bargaining relationships. In this section, we show first that conflating leaders and states as units of analysis can obscure reputation effects at both levels. Then, we highlight the consequences of using observational data to test unobservable implications of theoretical models. These claims motivate a research design that (a) accounts for the formation of reputations by both states and their individual leaders and (b) focuses on equilibrium relationships between leader and institutional change and dispute escalation in (c) theoretically informed samples.

National leaders make ultimate decisions over war and peace, but their freedom of action is often constrained by their institutional environment, making changes in both leaders and institutions relevant sources of private information. First, idiosyncratic differences render some leaders more hawkish or resolute than others (Horowitz, Stam, and Ellis 2015; Horowitz, McDermott, and Stam 2005; Palmer, London, and Regan 2004; Schultz 2005), and while some differences in background are observable to opponents (Yarhi-Milo et al. 2018), there remain many factors less observable *ex ante* to a leader's opponents. Second, different political arrangements can make the use of force more or less attractive to the same leader. By altering the size of political coalitions, the composition of relevant domestic audiences, or the rules by which leaders can secure support (or be punished), changes in political institutions constitute a state-level influence on a leader's resolve (Bueno de Mesquita et al. 2003; Croco 2011; Licht 2010, 2017; Saunders 2018; Weeks 2012; Weiss 2013), changes which can themselves create fresh informational traps.

Suppose that a hawkish incumbent leader oversees a transition from a relatively constrained institutional environment to an unconstrained one; her willingness to use force, previously limited by checks on her ability to display it, now interacts with a new set of domestic institutions to increase the state's effective level of resolve. Likewise, otherwise dovish incumbents might oversee institutional changes that bind them more closely to the preferences of a more hawkish constituency. Observing such an institutional transition, even as the incumbent remains in office, a foreign rival may expect a change in the leader's effective resolve. Lee Teng-hui, for example, remained president in Taiwan through major political reforms, including the direct election of the legislature in 1992. His subsequent statements about independence, viewed as provocative by Beijing and some members of his own party (Garver 1997, 90), helped throw into question his willingness to risk war with mainland China, setting the stage for the Taiwan Straits crises of the 1990s—crises



that may not have occurred had not institutional change portended changes in Lee's resolve over disputes with Beijing.

Most international relations scholarship analyzes the behavior of states with reference to the incentives of their leaders, but conflating units of analysis can pose inferential problems when public and private incentives diverge (Chiozza and Goe-mans 2004a, 604). How scholars answer the question of whether reputations inhere in governments or their leaders determines their interpretation of any particular crisis. If *political regimes* or states earn reputations, then Khrushchev's initiation of the Berlin Crisis in 1961 is surely a puzzle; the American commitment to the status quo in Berlin had been demonstrated in 1958. Yet, if *leaders* earn reputations, then the occurrence of the crisis makes sense: Khrushchev had incentives to test a new American president, while Kennedy had his own incentives to demonstrate hawkishness, resulting in a heightened risk of war. On the other hand, tensions between China and Taiwan in the mid-1990s make more sense when changes in *political regimes*, rather than in *leaders*, create informational traps.

The typical approach to looking for evidence of reputations is to search for links between past behavior and the occurrence or outcomes of present crises—that is, to look for the payment of reputation costs (see, inter alia, Gibler 2008; Huth and Russett 1993; Lupton 2018; Press 2005; Signorino and Tarar 2006; Weisiger and Yarhi-Milo 2015). Evidence that states pay reputation costs, in the form of changed opponent beliefs and subsequent challenges or unfavorable crisis outcomes, is inconsistent. Some states that show weakness fail to see future challenges, and their opponents do not seem to change their beliefs substantially in response to demonstrated strength or weakness (see, e.g., Hopf 1994; Mercer 1996; Press 2005). But these studies are predicated on observing an unobservable implication.<sup>10</sup>

Yet the apparent absence of large reputation costs in observational data is *consistent* with reputation theories. Situations in which reputation costs are highest are precisely those in which actors will work hardest to avoid them (see also Dafoe, Renshon, and Huth 2014, 385), showing dovishness only when they believe that doing so will not lead to significant consequences. If this relationship and the costs supporting it are systematically unobservable, then the search for reputation costs and their association with observable actions should be difficult, and inferences drawn from any such search are problematic.<sup>11</sup> Reputation theories *predict* no obvious observational link on average between signals of dovishness and the subsequent beliefs or behavior of opponents. Since their absence is also consistent with claims that reputations are irrelevant, finding no or limited reputation costs in the historical-empirical record provides no information.

Some studies that look for reputation costs in observational data nonetheless find them. Gibler (2008) shows that failing to honor alliance commitments reduces the chances that national leaders form future alliances, while Weisiger and Yarhi-Milo (2015) show that backing down in past militarized disputes increases the chances that states are targeted in future disputes. The past-actions research design remains subject to bias due to strategic censoring, in this case running the risk of

underestimating the role of reputation even if the expected relationship is recovered and statistically discernible. This approach makes tests conservative, but the attenuation of coefficients on reputation variables caused by strategic censoring makes comparisons across units of analysis—say, states and leaders—difficult (see Weisiger and Yarhi-Milo 2015). Our research design addresses this problem by focusing on empirical implications that are observable across all units in the sample, regardless of the information revealed by early interactions with opponents. This reduces the risks that differences across units of analysis are obscured by strategically biased samples. Experimental studies, on the other hand, can assess off-path behavior by testing predictions over how the public or a subset of elites attribute reputations in hypothetical situations (Renshon, Dafoe, and Huth 2018), but they may struggle to replicate the high-stress, high-stakes environment of international crises (Wilkenfeld and Brecher 2010). Even when elite samples are available (see Hafner-Burton et al. 2014), the unique social role of “national leader” can complicate inferences drawn from samples of individuals speculating about their likely choices in crisis situations. The best way to judge the empirical relevance of reputations is to leverage both experimental and observational data across different studies.

By contrast, attempts to avoid the problem of strategic censoring have highlighted more easily observable theoretical implications, uncovering further indirect evidence of how reputations form and affect behavior.<sup>12</sup> Walter (2006) shows that governments are less likely to accommodate separatists with multiple other groups sitting atop valuable land look for their own opportunities to secede, and Clare and Danilovic (2010, 2012) find evidence of similar “proactive” reputation-building, where states initiate or escalate disputes after having backed down in the past, especially when multiple potential opponents are watching and when the stakes are salient.<sup>13</sup> Crescenzi (2007) finds that states end up in military conflict against states that have fought states like them in the past,<sup>14</sup> and Dafoe (2012, chap. 4) finds that new leaders, who benefit most from cultivating a favorable reputation, are more likely to use force in militarized disputes, to experience longer disputes, to tolerate more casualties, and to win their disputes, than longer-serving leaders. Nonetheless, Walter (2006), Crescenzi (2007), and Clare and Danilovic (2010, 2012) estimate state-centric models, and Dafoe’s (2012) models are leader-centric. Each approach risks omitted-variable bias if reputations can form at both leader and state levels.

Finally, Rider (2013) examines Wolford’s (2007) hypothesis of reputation cycles driven by leadership turnover in the context of conventional arms races, which can act as a signal of resolve for leaders attempting to demonstrate their willingness to divert resources toward military spending. Competitive arms buildups are most common in the years immediately following leadership change in a dyad, becoming less likely as time passes. This prediction does not depend on the observability of reputation costs, only the potential for them to exist should actors *not* take the actions and draw inferences as predicted by the theory. Like Spaniel and Smith (2015) in their study of the imposition of economic sanctions and leader tenure,

we follow Rider's (2013) approach in our models of dispute escalation, though we privilege neither leadership turnover nor changes in political institutions as the sole source of new private information. Rather, leaders and political regimes can both be sources of variation in unobservable preferences like resolve and, as a consequence, the informational traps that generate cycles of reputation-building.

## Research Design

Finding evidence of reputations requires that we look for observable implications in observational data, and the historical–empirical record provides many instances of national leaders making decisions about dispute escalation. For each proposed informational trap, we expect that a recent change—from one leader to the next, from one set of political institutions to the next—will be independently associated with a probability of dispute escalation that begins high and then declines over time. But, as indicated by our theory, this pattern should be most likely to obtain in pairs of states with higher probabilities of facing one another in multiple disputes over time. Our two hypotheses follow directly:

**Hypothesis 1:** When future disputes are sufficiently likely, the probability that disputes escalate decreases in the time since the accession of a new national leader.

**Hypothesis 2:** When future disputes are sufficiently likely, the probability that disputes escalate decreases in the time since a change in political institutions.

The states in each of our samples are members of the international system according to the Correlates of War (COW 2017) data, and we identify national leaders with Archigos (Goemans, Gleditsch, and Chiozza 2009) for the main analyses and the Leader Experiences, Attributes, and Decisions (LEAD) data (Ellis, Horowitz, and Stam 2015) in the Online Appendix; we identify the beginning and end of political regimes with data from the Polity IV project (Marshall and Jaggers 2009). Our theoretical model also points us to several candidates for relevant samples in which states have common expectations of future disputes. A sample of all dyads in the international system, for example, contains so many states without expectations of disputes—much less repeated ones—that it would risk aggregation bias.<sup>15</sup> We first sample on pairs of strategic rivals, which view one another as long-term threats and account for a disproportionate amount of observed international conflict (Colaresi, Rasler, and Thompson 2008). Rivalries are an attractive sample due to their “strong issue disagreements, mutual suspicion, and *repeated militarized conflict*” (Hensel 1998, 66, emphasis added). They not only care about reputations but also have multiple opportunities to act on them, so we focus on Thompson's (2001) class of “strategic rivalries,” identified in historical sources using indicators of “when and with whom decision-makers thought they were in rivalry relationships” (p. 559). In strategic rivalries,

Actors interpret the intentions of others based on *earlier* behavior and forecasts about the *future* behavior of these other actors. The interpretation of these intentions leads to expectations about the likelihood of conflicts escalating to physical attacks. Strategic rivals anticipate some positive probability of an attack from their competitors over issues in contention. One side's expectations influence their own subsequent behavior toward their adversary and the process continues from there. (pp. 561-62, emphasis added)

The most recent version of the data identifies 165 dyadic rivalries between 1816 and 2010 (Thompson and Dreyer 2012). Taking into account data availability, our rivals sample includes 157 rival dyads from 1855 to 2007. During our sample period, rival dyads account for nearly 40 percent of all militarized interstate disputes (MIDs; Ghosn, Palmer, and Bremer 2004), averaging six MIDs over the sample period—substantially more than nonrival dyads.

Pairs of rivals are easy cases; if we are to see reputations “matter” anywhere, it should be in these dyads. However, rivals may not be the only sample in which reputation matters. Neighbors experience significantly more disputes than nonneighbors (Vasquez 1995), and major powers are both able and often willing to project power beyond their borders. Therefore, we also sample on land-contiguous (Stinnett et al. 2002) and politically relevant (Lemke and Reed 2001) dyads, where the latter sample includes pairs of land-contiguous neighbors and all dyads involving major powers as identified by the COW (2017) project. We expect evidence of reputations to be weaker in these two samples than in the sample of rival dyads, though stronger (if present) in each subsample than in the all-dyads sample.

We choose our unit of observation based on the time-series nature of the data and two theoretical loci of reputations. First, since informational traps imply that either side may take the steps that ultimately lead a dispute to escalate, we begin with the nondirected dyad-year. Second, to capture new private information at the levels both of leaders and states, we disaggregate the dyad-level unit of analysis into nondirected leader–regime dyad-years.<sup>16</sup> A new observation is created for a change of leadership *or* of regime type within a given dyad-year, such that there may be multiple observations within a given dyad-year. For example, there are two observations of the US–Soviet Union dyad in 1961 because Kennedy came to office on January 20, 1961, Khrushchev was in office the entire year, and there was no regime change in either state. The first observation starts from the first day of 1961 to January 20, and the second observation covers the remainder of the year. If new private information appears at multiple levels within a given dyad-year (e.g., changes both of leaders and of regime types), we add more observations accordingly. Our unit of analysis is appropriate for two reasons. First, it provides a more accurate value of our theoretical variables—a count of days since leader/regime changes. Second, and more importantly, we can match the occurrence of disputes with the exact period of observation in which new private information is introduced in a given year.

Next, our dichotomous outcome variable measures the escalation of an MID to the use of force. The errors across stages of dispute onset and escalation may be

correlated (cf. Reed 2000), so we use an estimator that models both dispute onset and escalation. At the selection stage, *Dispute onset* equals 1 if an MID occurs within a given observation. Since there may be multiple observations in a given dyad-year, we use the start date of MIDs identified for each observation to specify whether an MID occurred within that observation, as opposed to a previous or subsequent observation in the same year. Returning to the 1961 US-USSR example, a dispute began on July 25, so *Dispute onset* equals 0 in the first observation (January 1, 1961 to January 20, 1961) and 1 in the second observation (January 20, 1961 to December 31, 1961). If an MID has no clear start date, we code it as beginning on the first day of the month specified. If more than one MID breaks out during the observation, we focus on the MID with the highest hostility level.

After identifying a sample of disputes, we then look for escalatory behavior. Scholarship varies in how it operationalizes dispute escalation (see Braithwaite and Lemke 2011) and in the extent to which it distinguishes escalation from dispute onset. To match our theory as closely as possible, we assume that escalation involves the use of force, though it need not be mutual. Costly behaviors can demonstrate a side's own resolve or force the other to reveal its willingness to fight, but neither verbal threats nor simple displays of force constitute similarly costly behavior.<sup>17</sup> Our theory does not require that escalation be reciprocated because one side can reveal dovishness by refusing to escalate, which should also lead to a subsequent reduction in the probability of escalation. Therefore, *Dispute escalation* equals 1 when *at least* one side uses force, as defined in the MID data (i.e., a value of at least four on the hostility scale), and 0 otherwise.<sup>18</sup>

Finally, we model the onset of disputes because dispute dyads may differ in systematic, unmeasured ways from those that experience no disputes in the first place, leading to a correlation in the error terms across stages and potentially biased inferences. We use Heckman's (1979) model, adjusted for binary outcomes, to address the selection problem. The selection equation estimates the probability of dispute onset in the dyad, which we expect to be a function of time since leadership and regime changes. The outcome stage estimates the probability of dispute escalation among dispute dyads with

$$y_i^* = \beta_{0i} + \beta_1 \left( \text{days since leader change (log)}_i \right) + \beta_2 \left( \text{days since regime change (log)}_i \right) + \beta_k \mathbf{X}_i + \varepsilon_i,$$

where escalation is observed (i.e.,  $y_i = 1$ ) only when  $y_i^* > 0$ .  $\mathbf{X}_i$  is a vector of control variables, and standard errors are clustered by the dyad. We expect our theoretical variables to be negatively associated with escalation; that is, the probability of escalation is expected to be highest immediately after the introduction of new private information and then to decline over time. Therefore, the signs of both  $\beta_1$  and  $\beta_2$  should be negative.

Our two theoretical variables are (1) time since the accession of a new national leader and (2) time since a change in domestic institutions. For a given dyad, we

**Table 1.** Sources of New Private Information.

	State A Only	State B Only	Both States	Total
Leader only	1,259 (40.2%)	1,230 (39.27%)	11 (0.35%)	2,500 (79.82%)
Regime type only	154 (4.92%)	175 (5.59%)	3 (0.1%)	332 (10.61%)
Both changes	132 (4.21%)	131 (4.18%)	37 (1.18%) <sup>a</sup>	300 (9.57%)
Total	1,545 (49.33%)	1,536 (49.04%)	51 (1.63%)	3,132 (100%)

<sup>a</sup>Includes the observations in which both changes occur in at least one state.

calculate the time from a change (e.g., the date when a new leader comes into office) to the end date of a given observation for both states. Then, we use a weak-link specification (see Rider 2013, 585) to identify the side with the most recent change in the relevant locus of reputation. A change for either side can create an informational trap, and the more recently emerged trap should exert greater influence on the observation.<sup>19</sup>

The beginning of a leader's tenure is the start date when she comes into office according to Archigos (Goemans, Gleditsch, and Chiozza 2009).<sup>20</sup> For each dyad, we compute the tenure of each state's leader in days from the beginning of her time in office until the end date of the observation, then keep the shorter duration, which yields the variable *days since leader change*. Returning to the previous example, by the end of the second observation in 1961, Kennedy has been in office for 345 days and Khrushchev for 3208; we keep the former as the value of the shortest tenured leader's time in office in the observation.

Next, regime change is defined as a change from one regime type to another between time  $t - 1$  and time  $t$ . We adopt the Polity IV project's categories, which includes democracies (+6 to +10), anocracies (−5 to +5), and autocracies (−10 to −6) (Marshall and Jaggers 2009). In addition, there are three categories of political transition in which a state experiences foreign interruption (−66), interregnum or anarchy (−77), or transition (−88), each of which can introduce substantial uncertainty over a state's overall willingness to use force. The Polity IV data set records the beginning date of a new regime type based on these categories, which we label *days since regime change*.<sup>21</sup> Finally, since we expect that more information should be revealed in early as opposed to later years, we take the natural log of both measures of time since leader or regime change, creating  $\ln(\text{days since leader change} + 1)$  and  $\ln(\text{days since regime change} + 1)$ .<sup>22</sup>

Leaders and domestic institutions can change singly, making their effects easy to parse, or simultaneously, which sets a higher inferential hurdle. The problem of multicollinearity is particularly severe in the analysis of a small number of observations (e.g., Press 2005) as well as in studies with a strategically censored set of instances of demonstrated dovishness (e.g., Gibler 2008; Weisiger and Yarhi-Milo 2015). Table 1 provides descriptive statistics in the rivalry sample over the instances where states experience leader change only (e.g., from Eisenhower to Kennedy in the United

States), regime change only (e.g., Taiwan's transition to democracy as the head of state, Lee Teng-hui, remained the same), and changes in both leader and regime type. Leadership turnover accounts for roughly 80 percent of observed changes in private information, and changes in both leader and political institutions account for about 10 percent of new informational traps—roughly the same percentage of observations in which regime types change while the incumbent leader retains office. The data are weighted toward isolated instances of leadership turnover, but sufficient instances of changes in only regime type occur to facilitate a statistical comparison of the independent effects of both types of informational trap on dispute escalation.

We also include several control variables to account for potential confounding factors. First, pairs of democracies are uniquely peaceful, and throughout the sample period, they tend to be collectively aligned with a hegemonic power (McDonald 2015), creating stable clusters (Gleditsch 2002) of peacefully aligned states at relatively low risk of suffering regime change due to war. Thus, we control for the extent to which dyads are both democratic, using the Polity IV data set to construct *joint democracy* (Marshall and Jaggers 2009). A conventional approach is to code 1 if the Polity 2 score is greater than or equal to six for both states in a dyad and 0 otherwise. In our sample of strategic rivals, such a dichotomous measure would produce zero in nearly 95 percent of observations. Following Clare and Danilovic (2010), we create a continuous measure of *joint democracy*,

$$\text{Joint democracy} = \frac{\sqrt{(\text{polity2}_A + 11)(\text{polity2}_B + 11)}}{21},$$

which takes on higher values when both states approach higher democracy scores.<sup>23</sup> We also include a measure of the Polity democracy score for the shorter-lived regime, *Democracy*. Once democratization is complete, democracies tend to be long-lived (Boix and Stokes 2003; Przeworski and Limongi 1997), and democracies' conflict behavior can change over time, from early propensities for violence to later proclivities for peace (Mansfield and Snyder 2005), as a function of a process very different than reputation formation.

A leader's domestic political security is also related to both time in office and her conflict behavior. Chiozza and Goemans (2003, 2004b) find that the probability of crisis onset decreases with a leader's risk of losing office, though the opposite may be true in rivalries where leaders fear the consequences of over-cooperation (Colaresi 2004). If time in office is also associated with a leader's expectations about remaining in office (Chiozza and Goemans 2004a), we need a measure of political security to make sure our variable of interest captures reputational incentives rather than others such as diversionary incentives. We adopt Young's (2008) measure of job insecurity, an estimated probability that a leader will be removed from office based on the leader's age, the number of previous irregular leadership changes, and the level of democracy of the state. Given the temporal limitations of existing data, we follow Young's (2008) approach and create our estimate for each leader using

data from Goemans, Gleditsch, and Chiozza (2009) and Marshall and Jaggers (2009). Consistent with our measure of leader tenure, the resultant measure of *job insecurity* keeps the value of the leader with shorter tenure. The variable ranges from 0 to 1 (highest level of insecurity).<sup>24</sup>

We also include the *age* and the *previous times in office* for the leader with shorter tenure. Age is associated with an increased risk of losing office (Chiozza and Goemans 2004a) as well as an increased risk of conflict participation in democratic states (Horowitz, Stam, and Ellis 2015, 143-44). *Previous times in office* is associated with decreased risks of losing office (Chiozza and Goemans 2004a) and conflict participation (Chiozza and Goemans 2011), perhaps because leaders returning to office after a time of out of power—like Grover Cleveland in the United States—have more well-defined reputations, that is less uncertainty over their resolve, than truly new leaders.

Finally, we include several variables in the selection stage but exclude them from the outcome stage for identification purposes. First, *other rivals* counts the number of other strategic rivals that the shorter-tenured leader has in a given observation. As the number of rivals increases, disputes may be more likely to occur, but whether a larger number of rivals increases or decreases the chances of escalation is indeterminate; leaders may be cautious lest they waste resources that could be used against other enemies (see Treisman 2004), or they may be more inclined to fight to send a message to other rivals (see Walter 2006), but the direction of the effect isn't determined by the mere count of rivals. Second, the (logged) *distance* between state capitals (Gleditsch and Ward 2001) is generally associated with fewer opportunities for disputes to break out, though whether they escalate to the use of force may have more to do with where the military happens to be deployed and the territory between capital and frontier than simple distance.

*Peace duration* calculates the duration from the end date of the most recent MID to that of the given observation and converts it to a count of years. If the end day of an MID is unclear, we assume that such dispute ends on the last day of the month specified. If a dyad experiences an ongoing dispute, the variable is coded 0. To address the problem of temporal dependence, we include a cubic polynomial expansion of *peace duration* (see Carter and Signorino 2010) in the selection stage, though we exclude it from the outcome stage in order to identify the model. The duration dependence controls are inconsistently and only weakly related to escalation, while they have the necessary strong relationship with dispute onset (cf. Braithwaite and Lemke 2011; Reed 2000). Table 2 reports summary statistics for all the variables discussed above.

## Empirical Results

Table 3 reports the results of probit models uncorrected for the selection problem and Heckman probit models in the sample of strategic rivalry dyads. Models 1 and 2 measure regime change as a change from one type to another using Polity's three



**Table 2.** Summary Statistics.

Variables	Mean	SD	Minimum	Maximum	Observations
Dispute onset	0.111	0.315	0	1	8,466
Dispute escalation	0.713	0.452	0	1	944
Days since leader change	1,118	1,524	0	11,927	8,804
Days since regime change	5,596	6,655	0	40,827	8,804
Days since leader change (log)	6.150	1.518	0	9.387	8,804
Days since regime change (log)	7.8	1.532	0	10.617	8,804
Joint democracy	0.404	0.221	0.048	1	8,782
Democracy	0.199	0.399	0	1	8,804
Job insecurity	0.098	0.078	0.000	0.558	8,746
Age	54.751	12.014	11.94	88.4	8,775
Other rivalries	1.82	1.497	0	7	8,804
Distance	1,106.375	1,409.524	44	7,064	8,462
Peace duration	10.998	17.038	0	110	8,801

Note: SD = standard deviation.

categories, while models 3 and 4 code regime change as at least a three-point Polity change in any direction. Coefficients and statistical discernibility on our theoretical variables are similar—and consistent with our hypothesis—across all four models. Errors across the selection and outcome stages are negatively correlated ( $\rho < 0$ ), such that unobserved factors leading to dispute onset depress the probability of escalation, which is consistent with asymmetric-information accounts of deterrence and crisis bargaining (Fearon 1994; Reed 2000). Finally, we do not discuss the signs and significance of the coefficients on our control variables because they lack any controls for their own potential confounding factors.

All four models show that *days since leader change* is associated with a decrease in the probability of escalation. When two states engage in a dispute, that dispute is more likely to escalate if leadership turnover in the dyad has been relatively recent. Whether initial interactions reveal hawkishness or dovishness, the probability of dispute escalation declines over time (Hypothesis 1), with the period immediately following leadership change posing the greatest risk of dispute escalation. The coefficient is negative whether or not we model selection into disputes; it is also statistically discernible at  $p < .01$  in the uncensored samples of models 2 and 4 and at  $p < .05$  in the censored samples of models 1 and 3. Finally, *days since leader change* is associated with an increase in the probability of dispute onset but recall that reputations can produce any relationship between the passage of time and the onset of disputes, so despite its statistical discernibility, we can't learn anything from either this coefficient or the nondiscernible coefficient on *days since regime change*.

Our second theoretical variable, *days since regime change*, yields results consistent with Hypothesis 2. An increase in the time since a major change in political institutions appears unrelated to dispute onset, but once a dispute occurs, it is more

**Table 3.** Bootstrapped Estimation of Dispute Escalation between Strategic Rivals.

	Model 1		Model 2		Model 3		Model 4	
	Categorical Regime Change		Regime Change		Three-point Polity Change		Polity Change	
Dispute escalation								
Days since leader change	-0.143**	(.065)	-0.195***	(0.060)	-0.151**	(.061)	-0.195***	(.058)
Days since regime change	-0.106**	(.042)	-0.083**	(0.040)	-0.102***	(.040)	-0.087**	(.036)
Joint democracy	-0.386	(.352)	-0.404	(0.309)	-0.478	(.349)	-0.472	(.313)
Democracy	-0.315*	(.176)	-0.263	(0.182)	-0.322*	(.165)	-0.287*	(.173)
Job insecurity	1.555*	(.794)	1.294*	(0.752)	1.639***	(.622)	1.318**	(.625)
Age	0.004	(.004)	0.002	(0.003)	0.004	(.004)	0.003	(.003)
Previous times in office	-0.286***	(.103)	-0.225**	(0.099)	-0.291***	(.101)	-0.236**	(.099)
Constant	2.211***	(.557)	3.216***	(0.556)	2.219***	(.533)	3.189***	(.558)
Dispute onset								
Days since leader change			0.189***	(0.023)			0.180***	(.023)
Days since regime change			-0.016	(0.015)			-0.005	(.015)
Joint democracy			0.140	(0.199)			0.077	(.193)
Democracy			-0.048	(0.084)			-0.030	(.085)
Job insecurity			-0.240	(0.333)			-0.061	(.309)
Age			0.003	(0.002)			0.003*	(.002)
Previous times in office			-0.059	(0.041)			-0.058	(.041)
Other rivalries			-0.039*	(0.020)			-0.040**	(.020)
Distance			0.000**	(0.000)			0.000*	(.000)
Peace duration			-0.073***	(0.017)			-0.073***	(.017)
Peace duration <sup>2</sup>			0.002	(0.001)			0.001	(.001)
Peace duration <sup>3</sup>			-1.000	(1.000)			-0.000	(.000)
Constant			-2.160***	(0.197)			-2.188***	(.196)
$\rho$			-.600***				-.528***	
Model $\chi^2$	23.431		29.77		25.971		32.96	
Selection stage N			8,392				8,402	
Outcome stage N	936		934		937		934	

Note: Standard errors are in parentheses.

\* $p < .1$ .

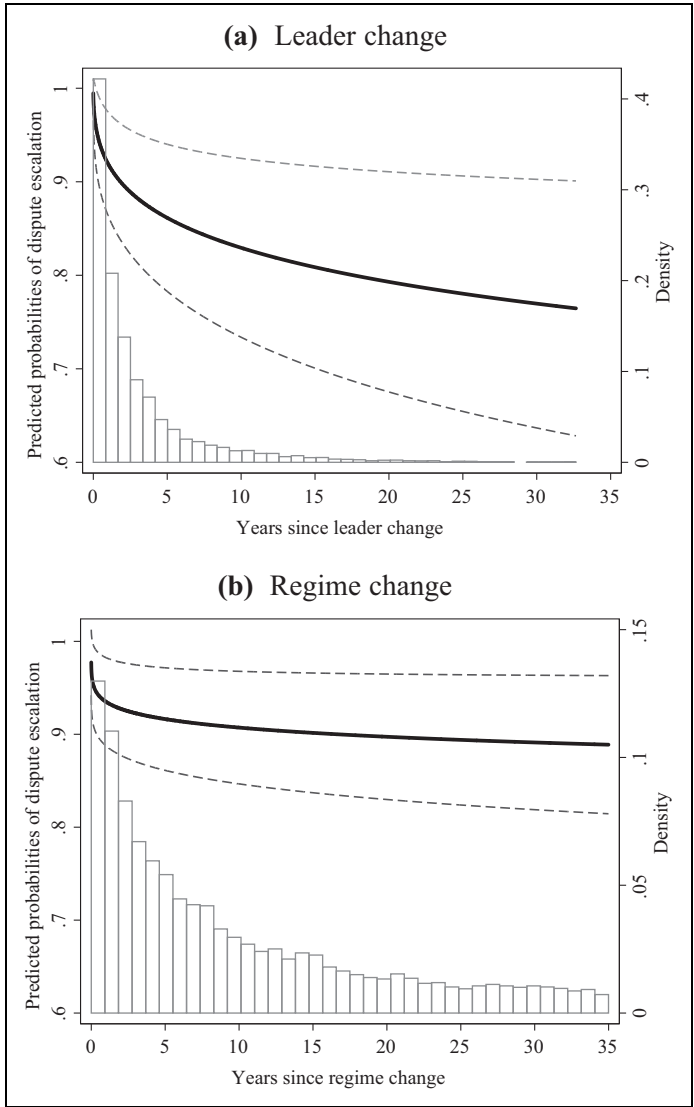
\*\* $p < .05$ .

\*\*\* $p < .01$ .

likely to escalate soon after a change in political institutions than later. This relationship is statistically discernible at  $p < .05$ . The coefficient on *days since regime change* in models 2 and 4 is smaller than that for *days since leader change (log)*, though their sizes are statistically indistinguishable ( $\chi^2_{(1)} = 2.17, p = .14$  in model 2;  $\chi^2_{(1)} = 2.63, p = .11$  in model 4), preventing us from making strong claims about the relative importance of leader- and state-specific reputations. Rather, the results are consistent with the idea that reputations can inhere in both national leaders *and* the states that they lead. Enduring questions about units and levels of analysis in international relations (Chiozza and Goemans 2011; Horowitz, Stam, and Ellis 2015; Singer 1961), particularly with respect to leaders or the states that they lead, may obscure the fact that reputations can be cultivated and attributed at *both* levels.

To get a substantive sense of the results, Figure 1 graphs predicted probabilities of escalation derived from model 2, where we de-log our variables and present results strictly in terms of time since changes in either leadership or political regimes (converted to yearly values to facilitate interpretation). The probabilities of escalation are quite high, as more than half of the disputes in our sample see some form of military force used. Figure 1(a) shows that the probability of escalation is highest immediately after leadership change before declining over time—substantially so through the first seven to eight years after a leader change. After about fifteen years, the probability of escalation becomes nearly stable. Figure 1(b) shows that the effects of time since regime change create a similar pattern, though the probability of escalation falls much more slowly than it does in Figure 1(a). The flattening of the curves in each case is also consistent with the intuition that the amount of information revealed is not constant across time. If conflict behaviors carry any information across crises or issues, they are most informative immediately after states acquire new leaders or new political institutions.

Strategic rivals share an expectation of future disputes by definition, but it's possible that reputation effects appear in wider samples as well. Table 4, then, begins with our rivalry sample and steadily expand the scope of our analysis to land-contiguous states (model 6), politically relevant dyads (model 7), and all dyads (model 8), effectively tracking decreases in the probability of future disputes ( $\delta$  in our theoretical model).<sup>25</sup> The results are largely consistent with our hypotheses. In models 5–7, *days since leader change* is associated with a decrease in the probability of dispute escalation and statistically discernible at  $p < .05$ , and the coefficient is largest in the rival sample and weakest in the sample of politically relevant dyads. We also find regime-related reputation effects, with similar magnitudes, in the samples of both contiguous dyads and politically relevant dyads. It's notable, though, that while *days since leader change* loses its statistical discernibility in the full sample (model 8), *days since regime change* retains both sign and discernibility ( $p < .05$ ). This may emerge from stronger regime-centered reputations in states that may interact less frequently than rivals or neighbors or from an unsatisfactory set of controls for *days since regime change*; we can't judge which



**Figure 1.** Predicted probabilities of dispute escalation.

of the two is true from this model. In summary, a comparison across different samples indicates that patterns of the cultivation and attribution of reputations exist most consistently between states with common expectations of future disputes, but how far these patterns travel outside of rivals, neighbors, and politically relevant dyads remains unclear.

**Table 4.** Estimating Dispute Escalation within Different Samples.

	Model 5		Model 6		Model 7		Model 8	
	Rivals		Contiguous		Politically Relevant		All	
Dispute escalation								
Days since leader change	-0.191**	(.074)	-0.165**	(.068)	-0.130**	(.062)	-0.068	(.058)
Days since regime change	-0.068**	(.034)	-0.048*	(.028)	-0.078***	(.028)	-0.080***	(.024)
Joint democracy	-0.359	(.364)	-0.482*	(.287)	-0.330	(.253)	-0.476**	(.230)
Democracy	-0.466***	(.156)	-0.157	(.153)	-0.305**	(.130)	-0.219*	(.115)
Job insecurity	1.392*	(.780)	0.109	(.807)	0.727	(.666)	0.533	(.619)
Age	0.008*	(.004)	0.005	(.004)	0.002	(.004)	0.004	(.003)
Previous times in office	-0.213**	(.102)	-0.126	(.079)	-0.104	(.073)	-0.092	(.070)
Constant	2.603***	(.689)	2.564***	(.528)	2.444***	(.521)	1.611***	(.480)
Dispute onset								
Days since leader change	0.184***	(.032)	0.134***	(.019)	0.109***	(.017)	0.075***	(.014)
Days since regime change	0.001	(.015)	-0.003	(.012)	0.007	(.012)	0.044**	(.012)
Joint democracy	0.233	(.189)	-0.086	(.141)	-0.146	(.108)	-0.006	(.079)
Democracy	0.035	(.090)	-0.108	(.077)	-0.146**	(.062)	-0.167***	(.048)
Job insecurity	0.187	(.522)	0.668**	(.307)	0.730***	(.232)	1.353***	(.187)
Age	-0.001	(.002)	-0.0002	(.0003)	-0.000	(.000)	-0.0002***	(.000)
Previous times in office	-0.008	(.043)	0.029	(.032)	0.021	(.028)	0.008	(.024)
Other rivalries	-0.037*	(.022)	0.003	(.017)	0.015	(.014)	-0.083***	(.012)
Distance	-0.009	(.012)	-0.013*	(.007)	-0.085***	(.007)	-0.151***	(.005)
Peace duration	-0.092***	(.024)	-0.109***	(.012)	-0.082***	(.007)	-0.084***	(.006)
Peace duration <sup>2</sup>	0.002	(.002)	0.003***	(.001)	0.002***	(.000)	0.002***	(.000)
Peace duration <sup>3</sup>	-0.000	(.000)	-0.000***	(.00)	-0.00***	(.00)	-0.00***	(.00)

(continued)

Table 4. (continued)

	Model 5		Model 6		Model 7		Model 8
	Rivals		Contiguous		Politically Relevant		All
Constant	-2.16***	(.224)	-1.933***	(.138)	-1.913***	(.112)	-2.214***
$\rho$	-.424***		-.396***		-.231***		-.013
Model $\chi^2$	31.433		24.69		29.637		31.417
Selection stage N	6,892		23,921		79,561		670,934
Outcome stage N	684		1,014		1,221		1,531

Note: Standard errors are in parentheses.

\* $p < .1$ .

\*\* $p < .05$ .

\*\*\* $p < .01$ .

## Conclusion

Fresh incentives to build reputations emerge after both leadership turnover and institutional change, and we show that the probability with which disputes escalate to the use of force is highest soon after new leaders take office and after major changes in political institutions. The risk then declines over time, as information is revealed before spiking again at the opening of a new informational trap. Our results therefore bridge the institutional and leader–attribute strands of research identified in the introduction to this special issue (Fuhrmann and Horowitz 2018): when leaders *or* the institutional constraints that bind them change, new informational traps open up that increase the risk that disputes escalate to the use of force. A leader's resolve is associated with some observable attributes such as military or rebel experiences (Horowitz et al. 2018), but our results highlight the importance of the unobservable factors that generate uncertainty over her resolve. And while we identify changes in leaders themselves and their institutional constraints only, future work may explore the reputational consequences of changes in advisers and figures surrounding national leaders (Saunders 2018) as well as the social coalitions that keep leaders in power (Licht 2010; Mattes, Leeds, and Matsumura 2016). The key is to look for the right evidence in the right place.

Our study also has implications for (a) testing hypotheses derived from informational theories of bargaining and war and (b) our understanding of the role of national leaders as units of analysis in international relations. Many attempts to uncover the effects of uncertainty on conflict behavior focus on situations in which private information is likely to be both present and significant (see, *inter alia*, Bennett and Stam 2004; Reed 2003; Slantchev 2004), but they vary in the extent to which they focus on another equally important part of the story: incentives to bluff, without which adversaries could simply state which deals they would accept in lieu of fighting, have their opponents believe them, and avoid war. A new leader in a state with no rivals, for example, has fewer incentives to display resolve by escalating crises than the leaders in our rivalry sample; in her case, we should not expect a strong relationship between time in office and dispute escalation, and estimating our model on a global sample bears this out. Analysts looking for evidence of a mechanism for war caused by informational problems should take care to measure not only conditions where private information over the attractiveness of war is present but also incentives to bluff about it. Early periods following leadership or regime change within rival or politically relevant dyads create just such an incentive to demonstrate resolve.

Finally, we provide evidence that the debate over units of analysis in international relations, particularly the distinction between leaders and the states in whose name they make decisions over war and peace (Bueno de Mesquita et al. 2003; Chiozza and Goemans 2004a; McGillivray and Smith 2008; Saunders 2011; Singer 1961), may be misplaced. Chiozza and Goemans (2011, chap. 6) show that leader-level variables account for a larger proportion of variance in empirical models of interstate conflict than state- or system-level variables, but we show that both individual

leaders and political institutions in which they operate make independent contributions to the processes of bargaining, reputation, and war. More concretely, looking for evidence of reputation-driven behavior at one level of analysis to the exclusion of the other can obscure, bias, or exaggerate the true underlying relationship. Levels of analysis may pose an inferential problem, but they need not be a theoretical problem. No particular level can claim precedence over the others, and scholars should focus on where and how their effects exist independently, and where and how their effects interact.

### **Authors' Note**

Previous versions of this article were presented at the 2015 International Studies Association Annual Convention, New Orleans, LA, and at the 2016 International Studies Association Annual Convention, Atlanta, GA.

### **Acknowledgments**

The authors would like to thank Phil Arena, Olga Chyzh, Sarah Croco, Scott Gartner, Kyo-suke Kikuta, Hans-Inge Langø, Amanda Licht, Carolina Moehlecke, Mark Nieman, Julie Phillips, Toby Rider, Emily Ritter, and Megan Thomas for helpful comments and suggestions. The authors thank ZhangMen Brewing and orgAsmo Curry in Taichung, Taiwan, and O'Dwyer's Irish Pub in Trier, Germany, for welcoming work environments and also thank Will Moore for the public impetus to (finally) test Hypothesis 1.


### **Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### **Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### **ORCID iD**

Cathy X. Wu  <http://orcid.org/0000-0001-7732-6066>

### **Supplemental Material**

Supplemental material for this article is available online.

### **Notes**

1. It's also high stakes, as prominent journalists specializing in political science research have bought into the narrative that reputational theories are "long debunked" (Fisher 2016) and that "no evidence" indicates that military action can convey relevant information to one's rivals (Fisher 2017).
2. Of course, if past behavior were factored into current estimates of power and interests, then it could have profound effects on present decision-making yet go unobserved



(Dafoe, Renshon, and Huth 2014; Weisiger and Yarhi-Milo 2015). This makes evidence of the attribution of reputations based on the diplomatic record problematic if one hopes to conduct a dispositive test—a problem that our empirical models are designed to solve.

3. On the uses (and abuses) of the term “resolve,” see Kertzer (2016).
4. However, we abstract away from the process by which different players develop different dispositions toward the use and cost of force (on this, see Kertzer 2016; Yarhi-Milo, Kertzer, and Renshon 2018).
5. Contrast this with Sartori (2005) and Guisinger and Smith (2002), who define reputations for honesty as simply *B*’s past record of behavior, which is perfectly known, as opposed to the more conventional definition we adopt, which involves *A*’s subjective assessment of some characteristic of *B*’s that may (or may not) be revealed by *B*’s actions. *B*’s actions in our theory may or may not reveal something about his underlying resolve rather than constituting its reputation.
6. This is a significant qualification to Schelling’s (1966) dictum about reputation or “face” being “worth fighting for.” If everyone fought for it, then fighting would convey no information, and “face” would be worthless. Pooling equilibria are uninformative.
7. Incentives to misrepresent private information are always present, but our argument is that reputations are a particular source of that incentive that change over time and emerge each time a new leader takes office or a new regime emerges in a state.
8. Both George and Smoke (1974) and Huth (1997, 1999) link this to the notion of interdependent commitments, but by focusing on the subjective costs of war, we identify a source of reputation that translates more easily across contexts than, say, promises of extended deterrence (see also Schelling 1966).
9. For an early statement of this problem, see Huth (1997).
10. Recent critiques of audience cost theories (Snyder and Borghard 2011) take the same approach.
11. Further, as shown by Treisman (2004), backing down today might even signal a future willingness to fight because it conserves resources. In this case, a reputation for weakness would not be inferred, but only because other factors—in this case, observable resource constraints—swamp the role of uncertainty.
12. Indirect evidence is no less evidentiary for being “indirect.” An appropriate analogy would be the search for evidence of gravity or of a comet responsible for a mass extinction millions of years in the past—each of which is inherently unobservable—but which imply a number of facts about the observable world that can be verified.
13. Our theoretical model reproduces this proactive reputation-building, where *B* initiates a crisis when he has a dovish reputation and room to change *A*’s mind. See Proposition 4.
14. Likewise, Crescenzi et al. (2012) find that states that have honored alliances in the past are more likely to form them in the present.
15. For a similar approach to distinguishing between information and commitment problems in disputes, see Moon and Souva (2016).
16. We recognize how complicated that label is, and we’re a little bit sorry about that. But only a little.

17. The “limited probes” discussed by George and Smoke (1974, 540-43, 569-72) are a special case of our measure of escalation, but it can also include levels of hostility up to full-scale war.
18. We follow Clare and Danilovic’s (2010) approach to escalation, which includes military action short of war. If we were to count only war and its 1,000 battle-death requirement as escalation, we would omit small-scale military operations conducted for reputational purposes. It is therefore consistent with our theoretical model to use a low threshold for escalation based on any use of force.
19. We use daily values to retain as much information as possible. We also convert the independent variables to yearly and monthly values and obtain similar results.
20. We use Archigos version 4.0, which covers the date of entry and exit of leaders from 1875 to 2010, and we reanalyze all models with Ellis, Horowitz, and Stam’s (2015) LEAD data in the Online Appendix.
21. We also measure regime change as a three-or-more point change (either increase or decrease) in the Polity 2 twenty-one point democracy–autocracy score at time  $t$  and time  $t - 1$ , and code *days since regime change* accordingly. The results are similar.
22. Bennett and Stam (2004, 180) use the age of dyads as a proxy for the revelation of information about preferences over time in their comparative test of a variety of conjectures about the causes of interstate conflict, but their analysis is cast at the state level and does not use a weak-link specification.
23. The Polity score is replaced with zero if a state is coded –66 (foreign interruption) or –77 (interregnum or anarchy) in the original Polity IV data set. We also test the model with the conventional dichotomous variable. The results are similar.
24. See Conrad and Ritter (2013) and Ritter (2014) for other uses of this variable in the context of international law and political violence.
25. Regime change is coded as at least three-point Polity change. To avoid too many nonevents in the larger samples, we exclude leaders who stay in office for less than thirty days, and the unit of analysis is nondirected leader–dyad-years. We still rely on the Archigos data set for our leader-related variables. To check the robustness of our analysis, we also code the leader-related variables using the LEAD data (Ellis, Horowitz, and Stam 2015) and present the results in Online Appendix. For the all-country-dyad data sets, we draw the militarized interstate disputes and LEAD variables from the NewGene software (Bennett, Poast, and Stam 2017).

## References

- Bennett, D. Scott, and Allan C. Stam. 2004. *The Behavioral Origins of War*. Ann Arbor: University of Michigan Press.
- Bennett, D. Scott, Paul Poast, and Allan C. Stam. 2017. “NewGene: A Conceptual Manual.” Accessed June 19, 2018. <https://ssrn.com/abstract=2596297>.
- Boix, Carles, and Susan C. Stokes. 2003. “Endogenous Democratization.” *World Politics* 55 (4): 517-49.
- Braithwaite, Alex, and Douglas Lemke. 2011. “Unpacking Escalation.” *Conflict Management and Peace Science* 28 (2): 111-23.

- Braumoeller, Bear F. 2012. *The Great Powers and the International System: Systemic Theory in Empirical Perspective*. New York: Cambridge University Press.
- Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge, MA: MIT Press.
- Carter, David B., and Curtis S. Signorino. 2010. "Back to the Future: Modeling Time Dependence in Binary Data." *Political Analysis* 18 (3): 271-92.
- Chiozza, Giacomo, and H. E. Goemans. 2003. "Peace through Insecurity: Tenure and International Conflict." *Journal of Conflict Resolution* 47 (4): 443-67.
- Chiozza, Giacomo, and H. E. Goemans. 2004a. "International Conflict and the Tenure of Leaders: Is War Still Ex Post Inefficient?" *American Journal of Political Science* 48 (3): 604-19.
- Chiozza, Giacomo, and H. E. Goemans. 2004b. "Avoiding Diversionary Targets." *Journal of Peace Research* 41 (4): 423-43.
- Chiozza, Giacomo, and H. E. Goemans. 2011. *Leaders and International Conflict*. Cambridge, MA: Cambridge University Press.
- Clare, Joe, and Vesna Danilovic. 2010. "Multiple Audiences and Reputation Building in International Conflicts." *Journal of Conflict Resolution* 54 (6): 860-82.
- Clare, Joe, and Vesna Danilovic. 2012. "Reputation for Resolve, Interests, and Conflict." *Conflict Management and Peace Science* 29 (1): 3-27.
- Clark, David H., and Timothy Nordstrom. 2005. "Democratic Variants and Democratic Variance: How Domestic Constraints Shape Interstate Conflict." *Journal of Politics* 67 (1): 250-70.
- Colaresi, Michael. 2004. "When Doves Cry: International Rivalry, Unreciprocated Cooperation, and Leadership Turnover." *American Journal of Political Science* 48 (3): 555-70.
- Colaresi, Michael, Karen Rasler, and William R. Thompson. 2008. *Strategic Rivalries in World Politics: Position, Space and Conflict Escalation*. Cambridge, MA: Cambridge University Press.
- Conrad, Courtenay R., and Emily Hencken Ritter. 2013. "Tenure, Treaties, and Torture: The Conflicting Domestic Effects of International Law." *Journal of Politics* 75 (2): 397-409.
- COW (Correlates of War). 2017. "State System Membership List, v2016." Online. Accessed June 19, 2018. <http://correlatesofwar.org/>.
- Crescenzi, Mark J. C. 2007. "Reputation and Interstate Conflict." *American Journal of Political Science* 51 (2): 382-96.
- Crescenzi, Mark J. C., Jacob D. Kathman, Katja Kleinberg, and Reed M. Wood. 2012. "Reliability, Reputation, and Alliance Formation." *International Studies Quarterly* 56 (2): 259-74.
- Croco, Sarah E. 2011. "The Decider's Dilemma: Leader Culpability, Domestic Politics, and War Termination." *American Political Science Review* 105 (3): 457-77.
- Dafoe, Allan. 2012. *Resolve, Reputation, and War: Cultures of Honor and Leaders' Time-in-office*. PhD thesis, University of California, Berkeley, CA.
- Dafoe, Allan, Jonathan Renshon, and Paul Huth. 2014. "Reputation and Status as Motives for War." *Annual Review of Political Science* 17:371-93.
- Danilovic, Vesna. 2002. *When the Stakes Are High: Deterrence and Conflict among Major Powers*. Ann Arbor: University of Michigan Press.

- Ellis, Cali Mortenson, Michael C. Horowitz, and Allan C. Stam. 2015. "Introducing the LEAD Data Set." *International Interactions* 41 (4): 718-41.
- Fearon, James D. 1994. "Signaling versus the Balance of Power and Interests." *Journal of Conflict Resolution* 38 (2): 236-69.
- Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49 (3): 379-414.
- Fearon, James D. 1997. "Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs." *Journal of Conflict Resolution* 41 (1): 68-90.
- Fisher, Max. 2016. "The Credibility Trap: Why Does a Long-debunked Theory Keep Leading the US into War?" Accessed June 19, 2018. <https://www.vox.com/2016/4/29/11431808/credibility-foreign-policy-war>.
- Fisher, Max. 2017. "Do U.S. Strikes Send a 'Message' to Rivals? There's No Evidence." *New York Times*, April 21. Accessed June 19, 2018. <https://www.nytimes.com/2017/04/21/world/do-us-strikes-send-a-message-to-rivals-theres-no-evidence.html>.
- Garver, John W. 1997. *Face off: China, the United States, and Taiwan's Democratization*. Seattle, WA: University of Washington Press.
- George, Alexander L., and Richard Smoke. 1974. *Deterrence in American Foreign Policy: Theory and Practice*. New York: Columbia University Press.
- 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: 133-54.
- Gibler, Douglas M. 2008. "The Costs of Reneging: Reputation and Alliance Formation." *Journal of Conflict Resolution* 52 (3): 426-54.
- Gleditsch, Kristian Skrede. 2002. *All International Politics Is Local: The Diffusion of Conflict, Integration, and Democratization*. Ann Arbor: University of Michigan Press.
- Gleditsch, Kristian Skrede, and Michael D. Ward. 2001. "Measuring Space: A Minimum Distance Database." *Journal of Peace Research* 38 (6): 749-68.
- Goemans, H. E., Kristian Skrede Gleditsch, and Giacomo Chiozza. 2009. "Introducing *Archigos*: A Data Set of Political Leaders." *Journal of Peace Research* 46 (2): 269-83. Accessed June 19, 2018. <http://mail.rochester.edu/hgoemans/data.htm>.
- Guisinger, Alexandra, and Alastair Smith. 2002. "Honest Threats: The Interaction of Reputation and Political Institutions in International Crises." *Journal of Conflict Resolution* 46 (2): 175-200.
- 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): 845-76.
- Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (1): 153-61.
- Hensel, Paul R. 1998. "Interstate Rivalry and the Study of Militarized Conflict." In *Conflict in World Politics: Advances in the Study of Crisis, War and Peace*, edited by Frank P. Harvey and Ben D. Mor, 162-204. New York: St. Martin's Press.
- Hopf, Ted. 1994. *Peripheral Visions: Deterrence Theory and American Foreign Policy in the Third World, 1965-1990*. Ann Arbor: University of Michigan Press.

- Horowitz and Fuhrmann (2018). "Leaders and Military Conflict." *Journal of Conflict Resolution* 62 (10): 2071-2085.
- Horowitz, Michael C., Allan C. Stam, and Cali M. Ellis. 2015. *Why Leaders Fight*. New York: Cambridge University Press.
- Horowitz, Michael, Rose McDermott, and Allan C. Stam. 2005. "Leader Age, Regime Type, and Violent International Relations." *Journal of Conflict Resolution* 49 (5): 661-85.
- Horowitz et al. (2018). "Sizing Up the Adversary: Leader Attributes, Credibility, and Reciprocation in International Conflict." *Journal of Conflict Resolution* 62 (10): 2179-2203.
- Huth, Paul K. 1997. "Reputations and Deterrence." *Security Studies* 7 (1): 72-99.
- Huth, Paul K. 1999. "Deterrence and International Conflict: Empirical Findings and Theoretical Debates." *Annual Review of Political Science* 2:25-48.
- Huth, Paul K., and Bruce Russett. 1993. "General Deterrence between Enduring Rivals: Testing Three Competing Models." *American Political Science Review* 87 (1): 61-73.
- Jervis, Robert. 1970. *The Logic of Images in International Relations*. Princeton, NJ: Princeton University Press.
- Kertzer, Joshua D. 2016. *Resolve in International Politics*. Princeton, NJ: Princeton University Press.
- Kreps, David M., and Robert Wilson. 1982. "Reputation and Imperfect Information." *Journal of Economic Theory* 27 (2): 253-79.
- Lemke, Douglas, and William Reed. 2001. "The Relevance of Politically Relevant Dyads." *Journal of Conflict Resolution* 45 (1): 126-44.
- Licht, Amanda A. 2010. "Coming into Money: The Impact of Foreign Aid on Leader Tenure." *Journal of Conflict Resolution* 54 (1): 58-87.
- Licht, Amanda A. 2017. "Hazards or Hassles: Modeling the Effect of Economic Sanctions on Leader Survival." *Political Science Research and Methods* 5 (1): 143-61.
- Lupton, Danielle L. 2018. "Signaling Resolve: Leaders, Reputations, and the Importance of Early Interactions." *International Interactions* 44 (1): 59-87.
- Mansfield, Edward D., and Jack Snyder. 2005. *Electing to Fight: Why Emerging Democracies Go to War*. Cambridge, MA: MIT Press.
- Marshall, Monty G., and Keith Jaggers. 2009. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2007." Polity IV Data set Version 2007. Accessed 30 July 2009. [www.systemicpeace.org/polity4](http://www.systemicpeace.org/polity4).
- Mattes, Michaela, Brett Ashley Leeds, and Naoko Matsumura. 2016. "Measuring Change in Source of Leader Support: The CHISOLS Dataset." *Journal of Peace Research* 53 (2): 259-67.
- McDonald, Patrick J. 2015. "Great Powers, Hierarchy, and Endogenous Regimes: Rethinking the Domestic Causes of Peace." *International Organization* 69 (3): 557-88.
- McGillivray, Fiona, and Alastair Smith. 2008. *Punishing the Prince: A Theory of Interstate Relations, Political Institutions, and Leader Change*. Princeton, NJ: Princeton University Press.
- Mearsheimer, John J. 1983. *Conventional Deterrence*. Ithaca, NY: Cornell University Press.
- Mercer, Jonathan. 1996. *Reputation and International Politics*. Ithaca, NY: Cornell University Press.

- Milgrom, Paul, and John Roberts. 1982. "Reputation, Predation, and Entry Deterrence." *Journal of Economic Theory* 27 (2): 280-312.
- Moon, Chungshik, and Mark Souva. 2016. "Audience Costs, Information, and Credible Commitment Problems." *Journal of Conflict Resolution* 60 (3): 434-58.
- Nalebuff, Barry. 1991. "Rational Deterrence in an Imperfect World." *World Politics* 43 (3): 315-35.
- Palmer, Glenn, Tamar R. London, and Patrick M. Regan. 2004. "What's Stopping You? The Sources of Political Constraints on International Conflict Behavior in Parliamentary Democracies." *International Interactions* 30 (1): 1-24.
- Powell, Robert. 1999. *In the Shadow of Power*. Princeton, NJ: Princeton University Press.
- Press, Daryl G. 2005. *Calculating Credibility: How Leaders Assess Military Threats*. Ithaca, NY: Cornell University Press.
- Przeworski, Adam, and Fernando Limongi. 1997. "Modernization: Theory and Facts." *World Politics* 49 (2): 155-83.
- Reed, William. 2000. "A Unified Statistical Model of Conflict Onset and Escalation." *American Journal of Political Science* 44 (1): 84-93.
- Reed, William. 2003. "Information, Power, and War." *American Political Science Review* 97 (4): 633-41.
- Renshon, Jonathan, Allan Dafoe, and Paul K. Huth. 2018. "Leader Influence and Reputation Formation in World Politics." *American Journal of Political Science* 62 (2): 325-39.
- Rider, Toby J. 2013. "Uncertainty, Salient Stakes, and the Causes of Conventional Arms Races." *International Studies Quarterly* 57 (3): 580-91.
- Ritter, Emily Hencken. 2014. "Policy Disputes, Political Survival, and the Onset and Severity of State Repression." *Journal of Conflict Resolution* 58 (2): 254-79.
- Russett, Bruce M. 1963. "The Calculus of Deterrence." *Journal of Conflict Resolution* 7 (2): 97-109.
- Sartori, Anne E. 2005. *Deterrence by Diplomacy*. Princeton, NJ: Princeton University Press.
- Saunders (2018). "Leaders, Advisers, and the Political Origins of Elite Support for War." *Journal of Conflict Resolution* 62 (10): 2117-2148.
- Saunders, Elizabeth N. 2011. *Leaders at War: How Presidents Shape Military Interventions*. Ithaca, NY: Cornell University Press.
- Schelling, Thomas C. 1966. *Arms and Influence*. New Haven, CT: Yale University Press.
- Schultz, Kenneth A. 2001. "Looking for Audience Costs." *Journal of Conflict Resolution* 45 (1): 32-60.
- Schultz, Kenneth A. 2005. "The Politics of Risking Peace: Do Hawks or Doves Deliver the Olive Branch?" *International Organization* 59 (1): 1-38.
- Signorino, Curtis S., and Ahmer Tarar. 2006. "A Unified Theory and Test of Extended Immediate Deterrence." *American Journal of Political Science* 50 (3): 586-605.
- Singer, J. David. 1961. "The Level-of-analysis Problem in International Relations." *World Politics* 14 (1): 77-92.
- Slantchev, Branislav. 2004. "How Initiators End Their Wars: The Duration of Warfare and the Terms of Peace." *American Journal of Political Science* 48 (4): 813-29.

- Snyder, Jack, and Erica D. Borghard. 2011. "The Cost of Empty Threats: A Penny, Not a Pound." *American Political Science Review* 105 (3): 437-56.
- Spaniel, William, and Bradley C. Smith. 2015. "Sanctions, Uncertainty, and Leader Tenure." *International Studies Quarterly* 59 (4): 735-49.
- Spence, Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics* 87 (3): 355-74.
- Stinnett, Douglas M., Jaroslav Tir, Philip Schafer, Paul F. Diehl, and Charles Gochman. 2002. "The Correlates of War Project Direct Contiguity Data, Version 3." *Conflict Management and Peace Science* 19 (2): 58-66.
- Thompson, William R. 2001. "Identifying Rivals and Rivalries in World Politics." *International Studies Quarterly* 45 (4): 557-86.
- Thompson, William R., and David Dreyer. 2012. *Handbook of International Rivalries: 1494-2010*. Washington, DC: Congressional Quarterly Press.
- Treisman, Daniel. 2004. "Rational Appeasement." *International Organization* 58 (2): 345-73.
- Vasquez, John A. 1995. "Why Do Neighbors Fight? Proximity, Interaction, or Territoriality." *Journal of Peace Research* 32 (3): 277-93.
- Walter, Barbara F. 2006. "Building Reputation: Why Governments Fight Some Separatists but Not Others." *American Journal of Political Science* 50 (2): 313-30.
- Weeks, Jessica L. 2012. "Strongmen and Straw Men: Authoritarian Regimes and the Initiation of International Conflict." *American Political Science Review* 106 (2): 326-47.
- Weisiger, Alex, and Keren Yarhi-Milo. 2015. "Revisiting Reputation: How Past Actions Matter in International Politics." *International Organization* 69 (02): 473-95.
- Weiss, Jessica Chen. 2013. "Authoritarian Signaling, Mass Audiences, and Nationalist Protest in China." *International Organization* 67 (1): 1-35.
- Wilkenfeld, Jonathan, and Michael Brecher. 2010. *International Crisis Behavior Project*. v.10 ed. Ann Arbor, MI: ICPSR Study #9286. Accessed February 7, 2011. <http://www.cidcm.umd.edu/icb/data>.
- Wolford, Scott. 2007. "The Turnover Trap: New Leaders, Reputation, and International Conflict." *American Journal of Political Science* 51 (4): 772-88.
- Yarhi-Milo et al. (2018). "Tying Hands, Sinking Costs, and Leader Attributes." *Journal of Conflict Resolution* 62 (10): 2149-2178.
- Young, Joseph K. 2008. *Repression, Dissent, and the Onset of Civil War*. Tallahassee: Florida State University. Dissertation.