Leaders, States, and Reputations: Appendix[*]

Updated July 11, 2020

Theoretical Model

We want to establish two claims: (a) that the probability of crisis escalation decreases over time regardless of whether a player reveals hawkishness or dovishness and (b) that the probability of crisis initiation can rise or fall as a function of what a player reveals about her type. To do this, we analyze a model that generalizes Wolford's (2007) game form, where players participate in two crises by construction—that is, crises are assumed to occur whether players would have initiated them or not to begin with.[1] In the present model, two players $i = \{A, B\}$ may bargain in two periods $t = \{1, 2\}$ over the division of a flow prize of unit value, where failure to reach an agreement in a period results in a war that divides the prize until the game ends or another opportunity bargain. These players can be states or their leaders, so to eliminate confusion we refer to them simply as "players," A female and B male, in our discussion of the model. B begins the game with private information over his costs for war, i.e. his resolve, and A must make decisions over whether and how much to demand in a revision of the status quo uncertain over just how much she can extract from B without provoking a war. We also analyze a version in which B can be the first mover and can choose whether to initiate a crisis in the first period. Repeated interactions allow for changes in reputations and analysis of their consequences, and by introducing an opportunity to initiate we can draw the critical distinction between the onset and escalation of disputes. Thus, we can establish all three of our claims with a minimum of moving parts.

The game begins with Nature drawing B's costs for war b from the uniform distribution $b \sim \mathrm{U}(\underline{b}, \overline{b})$ and revealing that information to B alone. A knows only the distribution from which B's type is drawn, which constitute her prior beliefs (the reputation with which B enters the game) at the beginning of the game. Then, depending on the variant of the model under study, a player tapped as the first mover (A or B) chooses whether or not to initiate a crisis. If the first mover passes, then the status quo division of the prize, $(q, 1-q)$ for A and B respectively, remains in place for $t = 1$ and the game moves time $t = 2$, where a crisis occurs with probability $\delta \in (0, 1)$ (otherwise, the game ends with both players receiving zero).

[*]We're grateful to Chen Wang for catching a typo on the bottom pf page 10.

[1]We do, however remove the endogenous leader change machinery from Wolford's (2007) model, as the results do not depend on it.

This ensures that, when they act in the first period, players must consider the possibility of future interactions in which reputations developed or information gained in the first period can shape their welfare. Whether they occur in the first or the second period, all crises are the same. The uninformed player (A) proposes a division of the prize, $(d, 1-d)$ where $d = \{x, y, w, z\}$ and a larger value of d allocates more to A and less to B, that the informed player (B) can accept or reject. Acceptance implements the proposed division of the prize for the current period, and rejection leads to a costly lottery (war) that allocates the whole prize to the winner for the current period only. If the crisis occurs in the first period, for example, the outcome of the war is up for renegotiation if a second-period crisis occurs. Therefore, the sequence of play is

1. N draws $b \sim \mathrm{U}(\underline{b}, \overline{b})$ and reveals it to B.
2. At $t = 1$, the initiator (A or B) initiates a crisis or passes.
3. If a crisis occurs, A proposes $(x, 1-x)$, and B accepts (peace) or rejects (war).
4. At $t = 2$, a crisis occurs with probability $\delta \in (0, 1)$.
5. If the initiator passed at $t = 1$, A proposes $(z, 1-z)$, and B accepts or rejects.
6. If war occurred at $t = 1$, A proposes $(w, 1-w)$, and B accepts or rejects.
7. If peace obtained at $t = 1$, A proposes $(y, 1-y)$, and B accepts or rejects.

Preferences are standard for crisis bargaining models (see, *inter alia*, Fearon 1995, Powell 1999), in that payoffs are linear in shares of the prize, e.g. $(x, 1-x)$ for a settlement at $t = 1$, and war is a costly lottery: A wins a war with probability p, B with probability $1-p$, and each player pays an upfront cost to participate (a and b, respectively).To keep threats credible and offers interior, we assume that $a < p$ and $\overline{b} < 1-p$. If, for example, the players strike a bargain at $t = 1$ and go to war at $t = 2$, payoffs are

$$EU_A = x + \delta(p - a) \quad \text{and} \quad EU_B = 1 - x + \delta(1 - p - b),$$

and if no crisis occurs at $t = 1$ before players strike a deal at $t = 2$, payoffs are

$$EU_A = q + \delta z \quad \text{and} \quad EU_B = 1 - q + \delta(1 - z),$$

where δ behaves like a standard shared discount factor in repeated games and represents the shadow of an uncertain future whose consequences may shape today's actions. The higher δ, the more likely are the players to confront one another again in the future—and since distributive outcomes do not travel, informational consequences do, allowing us to isolate the effect of reputation concerns on our outcomes of interest.

We analyze two variants of the model, one in which each player is the potential first mover, proving for each the existence of a semi-separating Perfect Bayesian Equilibrium (PBE) in which reputations form in the first period and shape the probability of war in the second. PBE requires that strategies are sequentially rational and consistent with beliefs updated via Bayes' Rule on the equilibrium path, and in a semi-separating PBE, groups of player-types take a the same action, allowing some belief updating to occur. In this section, we first work out the logic of, or the "why" behind, each equilibrium and showing that the link between past actions and present crisis behavior is weak. Then, we conduct comparative statics analysis to pin down expectations about equilibrium behavior. We show that (a) the probability of war declines over time on the equilibrium path and (b) whether the probability of crisis initiation rises or falls over time depends on the content of B's reputation.

The equilibria of interest share a fundamental strategic tension, differing only in which player has an opportunity to initiate a crisis in the first period. Crises entail a risk-return tradeoff (see Powell 1999, Ch. 3): the more the uninformed player demands, the better she does in a settlement if B accepts, but larger demands also entail a greater risk that B finds the proposal unacceptable and rejects. A's proposal in the first period (x) balances these priorities by tolerating a probability that B rejects. This separation of types, rejection by the resolute and acceptance by the irresolute, allows A to update her beliefs about B's costs for war ahead of the possible second crisis; in other words, she attributes a reputation to B that informs her subsequent bargaining strategy. But while A can force a revelation of private information, B has incentives to fight in the first period in order to convince A of her resolve and receive better bargains in the second. Both sides have incentives to take risks in the first period—A to initiate a crisis and then make demands that might be rejected—B to initiate a crisis and then fight a war, in the service of positioning themselves favorably for a potential future crisis. Doing otherwise either fails to force the revelation of useful information (for A) or to cultivate a favorable reputation (for B), and as such each player is trapped by its incentives to take actions that, under complete information, it would avoid.

Proposition 1 (A Initiates)**.** *The following strategies and beliefs constitute a Perfect Bayesian Equilibrium when A is the first mover. For A,*

- *At $t = 1$, initiate iff $q < q_A$; propose $x = x^*$.*
- *At $t = 2$, believe $b \in (\underline{b}, b_1)$ and propose $w = w^*$ if war at $t = 1$; believe $b \in [b_1, \overline{b})$ and propose $y = y^*$ if peace at $t = 1$; believe $b \in (\underline{b}, \overline{b})$ and propose $z = z^*$ if no crisis at $t = 1$.*

For B,

- *At $t = 1$, accept iff $b \geq b_1$.*
- *At $t = 2$, accept w^* iff $b \geq b_2^w$, accept y^* iff $b \geq b_2^y$, and accept z^* iff $b \geq b_2^z$.*

Proof of Proposition 1. PBE requires sequentially rational strategies and weak consistency of beliefs—i.e., that they are updated according to Bayes' Rule where possible. All actions are taken in equilibrium, so there are no out-of-equilibrium beliefs to specify. We identify a cut-point strategy for B that parses the type space with types indifferent between taking "higher" and "lower," i.e. more and less costly, actions, which allows for a truncation of A's beliefs into smaller uniform distributions.

To prove existence, we begin by defining a generic acceptance rule for B at time $t = 2$, then deriving sequentially rational strategies and consistent beliefs for all histories leading to $t = 2$. Letting $d = \{w, y, z\}$, B accepts iff

$$1 - d \geq 1 - p - b \Leftrightarrow b \geq d - p \tag{1}$$

and rejects otherwise. Next, if A passes at $t = 1$, she believes $b \sim \mathrm{U}(\underline{b}, \overline{b})$ at $t = 2$, so z maximizes

$$EU_A(z) = \int_{\underline{b}}^{z-p} (p-a)\,\mathrm{d}U\mathrm{d}b + \int_{z-p}^{\overline{b}} (z)\,\mathrm{d}U\mathrm{d}b \tag{2}$$

by solving

$$\frac{\partial EU_A(z)}{\partial z} = -a + 2p - 2z + \overline{b} = 0$$

to yield

$$z^* = p + \frac{\overline{b} - a}{2},$$

which we know to be a maximum since $\partial^2 EU_A(z)/\partial z^2 = -2$. If A initiates at $t = 1$ and B accepts x^*, A believes that $b \sim \mathrm{U}[b_1, \overline{b})$, and by Inequality (5) we know that B accepts iff $b \geq y - p$. So A maximizes

$$EU_A(y) = \int_{b_1}^{y-p} (p-a)\,\mathrm{d}U\mathrm{d}b + \int_{y-p}^{\overline{b}} (y)\,\mathrm{d}U\mathrm{d}b \tag{3}$$

by solving

$$-a2p - 2y + \overline{b} = 0$$

to yield

$$y^* = p + \frac{\overline{b} - a}{2},$$

which we know to be a maximum since $\partial^2 EU_A(y)/\partial y^2 = -2$. Finally, if A initiates at $t = 1$ and B rejects x^*, A believes that $b \sim \mathrm{U}(\underline{b}, b_1)$, and by Inequality (5) we know that B accepts iff $b \geq w - p$. So A maximizes

$$EU_A(w) = \int_{\underline{b}}^{w-p} (p-a)\,\mathrm{d}U\mathrm{d}b + \int_{w-p}^{b_1} (w)\,\mathrm{d}U\mathrm{d}b \tag{4}$$

by solving

$$-a2p - 2w + b_1 = 0$$

to yield

$$w^* = p + \frac{b_1 - a}{2},$$

which we know to be a maximum since $\partial^2 EU_A(w)/\partial y^2 = -2$.

Now consider B's acceptance rule at $t = 1$, which defines a type b_1 that is indifferent over accepting and rejecting A's proposal x^*. Acceptance yields $1-x$ at $t = 1$ and, since B believes $b \sim \text{U}[b_1, \overline{b})$ and proposes y^* in response, war at $t = 2$; rejection yields war at $t = 1$ and, since B believes $b \sim \text{U}(\underline{b}, b_1)$ and proposes w^* in response, $1 - w^*$ at $t = 2$. Therefore, b_1 solves

$$1 - x + \delta(1 - p - b_1) = 1 - p - b_1 + \delta(1 - w^*) \Leftrightarrow b_1 = \frac{2(x-p)+\delta a}{2-\delta},$$

such that lower-cost types reject and higher-cost types accept. Anticipating the effects on its prior beliefs, A maximizes

$$EU_A(x) = \int_{\underline{b}}^{w^*-p} (p - a + \delta(p-a))\,\mathrm{d}U\mathrm{d} + \int_{w^*-p}^{b_1} \left(p - a + \delta w^*\right)\mathrm{d}U\mathrm{d} +$$
$$\int_{b_1}^{y^*-p} (x + \delta(p-a))\,\mathrm{d}U\mathrm{d} + \int_{y^*-p}^{\overline{b}} \left(x + \delta y^*\right)\mathrm{d}U\mathrm{d}$$

by solving

$$\frac{\partial EU_A(x)}{\partial x} = \overline{b} + \frac{6(a-p+x)}{\delta-2} + \frac{4(a-p+x)}{(\delta-2)^2} + a = 0$$

to yield

$$x^* = p + \frac{1(\delta(2+\delta)-4) + \overline{b}(2-\delta)^2}{8-6\delta},$$

which we know to be a maximum since $\partial^2 EU_A(x)/\partial x^2 = -(8-6\delta)/(2-\delta)^2 < 0$.

Finally, we consider A's choice over passing or initiating. If she passes, the status quo remains in place for $t = 1$ but at $t = 2$ a crisis in which she retains her prior beliefs occurs with probability δ. A initiates when

$$EU_A(x^*) > q + \delta EU_A(z^*),$$

or when

$$q < -\frac{2\overline{b}(a(\delta(\delta+2)-4) - 6\delta p + 8p) - 4(3\delta-4)\underline{b}(a-p) + (\delta-2)^2\overline{b}^2 + a^2(\delta-2)^2}{4(3\delta-4)} = q_A$$

and passes otherwise.

Strategies are sequentially rational and beliefs are updated according to Bayes' Rule where possible, so the proposed PBE exists. □

Proposition 2 (B Initiates)**.** *The following strategies and beliefs constitute a Perfect Bayesian Equilibrium when B is the first mover. For A,*

- *At $t = 1$, propose $x = x^*$.*
- *At $t = 2$, believe $b \in (\underline{b}, b_1)$ and propose $w = w^*$ if war at $t = 1$; believe $b \in [b_1, b_0)$ and propose $y = y^*$ if peace at $t = 1$; believe $b \in [b_0, \overline{b})$ and propose $z = z^*$ if no crisis at $t = 1$.*

For B,

- *At* $t = 1$*, initiate iff* $q > q_B$*; accept iff* $b \geq b_1$.
- *At* $t = 2$*, accept* w^* *iff* $b \geq b_2^w$*, accept* y^* *iff* $b \geq b_2^y$*, and accept* z^* *iff* $b \geq b_2^z$.

Proof of Proposition 2. PBE requires sequentially rational strategies and weak consistency of beliefs—i.e., that they are updated according to Bayes' Rule where possible. All actions are taken in equilibrium, so there are no out-of-equilibrium beliefs to specify. We identify a cut-point strategy for B that parses the type space with types indifferent between taking "higher" and "lower," i.e. more and less costly, actions, which allows for a truncation of A's beliefs into smaller uniform distributions.

To prove existence, we begin by defining a generic acceptance rule for B at time $t = 2$, then deriving sequentially rational strategies and consistent beliefs for all histories leading to $t = 2$. Letting $d = \{w, y, z\}$, B accepts iff

$$1 - d \geq 1 - p - b \Leftrightarrow b \geq d - p \tag{5}$$

and rejects otherwise. Next, if B passes at $t = 1$, A believes $b \sim \text{U}(b_0, \overline{b})$ at $t = 2$, so z maximizes

$$EU_A(z) = \int_{b_0}^{z-p} (p - a)\,\mathrm{d}U\,\mathrm{d}b + \int_{z-p}^{\overline{b}} (z)\,\mathrm{d}U\,\mathrm{d}b \tag{6}$$

by solving

$$\frac{\partial EU_A(z)}{\partial z} = -a + 2p - 2z + \overline{b} = 0$$

to yield

$$z^* = p + \frac{\overline{b} - a}{2},$$

which we know to be a maximum since $\partial^2 EU_A(z)/\partial z^2 = -2$. If B initiates at $t = 1$ and accepts x^*, A believes that $b \sim \text{U}[b_1, b_0)$, and by Inequality (5) we know that B accepts iff $b \geq y - p$. So A maximizes

$$EU_A(y) = \int_{b_1}^{y-p} (p - a)\,\mathrm{d}U\,\mathrm{d}b + \int_{y-p}^{b_0} (y)\,\mathrm{d}U\,\mathrm{d}b \tag{7}$$

by solving

$$-a2p - 2y + b_0 = 0$$

to yield

$$y^* = p + \frac{b_0 - a}{2},$$

which we know to be a maximum since $\partial^2 EU_A(y)/\partial y^2 = -2$. Finally, if B initiates at $t = 1$ and rejects x^*, A believes that $b \sim \mathrm{U}(\underline{b}, b_1)$, and by Inequality (5) we know that B accepts iff $b \geq w - p$. So A maximizes

$$EU_A(w) = \int_{\underline{b}}^{w-p} (p-a)\,\mathrm{d}U\mathrm{d}b + \int_{w-p}^{b_1} (w)\,\mathrm{d}U\mathrm{d}b \tag{8}$$

by solving

$$-a2p - 2w + b_1 = 0$$

to yield

$$w^* = p + \frac{b_1 - a}{2},$$

which we know to be a maximum since $\partial^2 EU_A(w)/\partial y^2 = -2$.

Now consider B's acceptance rule at $t = 1$, which defines a type b_1 that is indifferent over accepting and rejecting A's proposal x^*. Acceptance yields $1 - x$ at $t = 1$ and, since B believes $b \sim \mathrm{U}[b_1, \overline{b})$ and proposes y^* in response, war at $t = 2$; rejection yields war at $t = 1$ and, since B believes $b \sim \mathrm{U}(\underline{b}, b_1)$ and proposes w^* in response, $1 - w^*$ at $t = 2$. Therefore, b_1 solves

$$1 - x + \delta(1 - p - b_1) = 1 - p - b_1 + \delta(1 - w^*) \Leftrightarrow b_1 = \frac{2(x-p) + \delta a}{2 - \delta},$$

such that lower-cost types reject and higher-cost types accept. Anticipating the effects on its prior beliefs, A maximizes

$$\begin{aligned} EU_A(x) = &\int_{\underline{b}}^{w^*-p} (p - a + \delta(p-a))\,\mathrm{d}U\mathrm{d} + \int_{w^*-p}^{b_1} \left(p - a + \delta w^*\right)\mathrm{d}U\mathrm{d} + \\ &\int_{b_1}^{y^*-p} (x + \delta(p-a))\,\mathrm{d}U\mathrm{d} + \int_{y^*-p}^{b_0} \left(x + \delta y^*\right)\mathrm{d}U\mathrm{d} \end{aligned}$$

by solving

$$\frac{\partial EU_A(x)}{\partial x} = \frac{a(\delta(\delta+2) - 4) + (\delta-2)^2 b_0 - 2(3\delta - 4)(p - x)}{(\delta - 2)^2} = 0$$

to yield

$$x^* = p + \frac{1(\delta(2+\delta) - 4) + b_0(2-\delta)^2}{8 - 6\delta},$$

which we know to be a maximum since $\partial^2 EU_A(x)/\partial x^2 = -(8 - 6\delta)/(2-\delta)^2 < 0$.

Finally, there exists a type b_0 that is indifferent over initiating, which entails accepting x^* at $t = 1$ and y^* at $t = 2$, and passing, which secures the status quo for $t = 1$ but allows A to update her beliefs to $b \sim \mathrm{U}[b_0, \overline{b})$ for the crisis at $t = 2$. Therefore, b_0 solves

$$1 - x^* + \delta\left(1 - y^*\right) = 1 - q + \delta\left(1 - z^*\right)$$

such that

$$b_0 = \frac{\delta(3\delta-4)\overline{b} + a(\delta(\delta+2)-4) - 2(3\delta-4)(p-q)}{2\left(\delta^2-2\right)}.$$

When $b < b_0$, B initiates, and when $b \geq b_0$, B passes.

Strategies are sequentially rational and beliefs are updated according to Bayes' Rule where possible, so the proposed PBE exists. □

Propositions 1 and 2 describe equilibria in which types of B parse themselves into groups by taking certain actions in equilibrium, where those groups are defined by a series of cut-points above and below which types take different actions. The higher the range a type falls in, the more dovish is B's reputation, but as b falls into lower ranges, the more hawkish its reputation. As we'll discuss further below, the primary difference across the two equilibria is that B only initiates a crisis of his own when sufficiently resolute, the act of initiating the crisis already partitions the type space, and thus A's beliefs, to $b \in (\underline{b}, b_0)$. Notably, types $b \in [b_0, \overline{b})$ reveal themselves as such, consigning themselves to a dovish reputation and a raw deal at $t = 2$. Such an outcome is too costly to avoid for two reasons; first, initiating a crisis means having to wait to reap the benefits of a better reputation, and second, if the most irresolute types were to pool with more resolute types, then A's beliefs wouldn't change sufficiently to produce any worthwhile gains in the future. But for the remaining, relatively resolute types of B, avoiding such a reputation is worth the cost. Unless otherwise noted, we write $b_h = \{\overline{b}, b_0\}$ for the highest type that enters the first crisis.

Provided that a crisis does occur at $t = 1$, in each equilibrium A makes an initial proposal x^* that B accepts when sufficiently irresolute ($b \geq b_1$) and rejects when resolute ($b < b_1$). This allows A to update her beliefs going into the second crisis, truncating the range of the distribution from $b \sim (\underline{b}, b_h)$ at the beginning of the crisis to $b \sim (\underline{b}, b_1)$ if B rejects and $b \sim [b_1, b_h)$ if B accepts. These updated beliefs constitute B's reputation for the second period, and if a crisis occurs A chooses her demand based on those new beliefs, demanding more than she did at $t = 1$ if B proves dovish and less if B proves hawkish. To the extent that reputation costs are paid in the second period, they manifest in tolerating worse deals in future crises (cf. Wolford 2007, p. 778), but not in a higher probability of war, because as A's beliefs are updated in light of new information she is able to fashion her demands so as to reduce (weakly) the risks of war, whether B reveals itself to be dovish or hawkish. As long as the shadow of the future looms long enough—i.e. as long as δ is sufficiently large, which we associate with conditions of strategic rivalry—then later crises should be less prone to escalate than earlier crises. We prove this claim for the model in which A initiates, but since the only difference in the two after the first crisis occurs is the value of b_h, the argument applies equally to the model in which B initiates.

Proposition 3 (The Probability of War)**.** *For δ sufficiently high and a crisis at $t = 1$, the probability of war falls from $t = 1$ to $t = 2$ on the equilibrium path.*

Proof of Proposition 3. Since b is distributed uniformly and B plays a cut-point strategy, the probability of war is the probability that B is of a type that rejects an equilibrium demand

d^*. In the first period,

$$\Pr(\text{War}_1) = \frac{b_1 - \underline{b}}{\overline{b} - \underline{b}},$$

and in the second,

$$\Pr(\text{War}_2^w) = \frac{w^* - p - \underline{b}}{b_1 - \underline{b}} \quad \text{and} \quad \Pr(\text{War}_2^y) = \frac{y^* - p - b_1}{\overline{b} - b_1},$$

where $w^* - y$ and $y^* - y$ are, respectively, the cut-points above which types b accept and below which they reject equilibrium demands w^* and y^*. First, we solve $\Pr(\text{War}_1) > \Pr(\text{War}_2^w)$ to show that war is less likely at $t = 2$ than at $t = 1$ if B rejected x^* at $t = 1$. This inequality is satisfied for any value of a when

$$\delta > \frac{2\left(a - \overline{b} + 2\underline{b}\right)}{2a - \overline{b} + 3\underline{b}}.$$

Second, we solve $\Pr(\text{War}_1) > \Pr(\text{War}_2^y)$ to show that war is less likely at $t = 2$ than at $t = 1$ if B accepted x^* at $t = 1$. This inequality is satisfied for any value of a when

$$\delta > -2\sqrt{2}\sqrt{\frac{\left(\overline{b} - \underline{b}\right)\left(\overline{b} + a\right)}{\left(\overline{b} - 9\underline{b} - 8a\right)^2}} + \frac{3\left(\overline{b} - \underline{b}\right)}{\overline{b} - 9\underline{b} - 8a} + 1.$$

Specifically, when $a \leq \overline{b} - 2\underline{b}$, both inequalities are satisfied for any value of δ, but when $a > \overline{b} - 2\underline{b}$, these conditions on δ are required. Therefore, when δ is sufficiently high, the probability of war falls from a crisis at $t = 1$ to a crisis at $t = 2$, regardless of the information revealed by the first crisis. □

Finally, we want to establish that the probability of crisis onset depends on the content of B's reputation and not simply the revelation of information before the crisis, as is the case for escalation in Proposition 3. Proposition 4 proves this claim by examining the conditions in each game under which the first mover initiates a crisis as a function of A's beliefs over B's type when the game begins—that is, as a function of the reputation with which B enters a given interaction.

Proposition 4 (Crisis Initiation)**.** *The probability of crisis initiation may increase or decrease as a function of revealed information. A's initiation constraint is harder to satisfy as $\underline{b}$ increases and easier to satisfy as $\overline{b}$ increases, while the probability that B initiates can increase or decrease in both $\underline{b}$ and $\overline{b}$.*

Proof of Proposition 4. Begin with the game in which A moves first. A initiates when $q < q_A$, where

$$q_A = -\frac{2\overline{b}(a(\delta(\delta+2)-4) - 6\delta p + 8p) - 4(3\delta-4)\underline{b}(a-p) + (\delta-2)^2\overline{b}^2 + a^2(\delta-2)^2}{4(3\delta-4)}.$$

The first partial derivative with respect to $\underline{b}$ is

$$\frac{\partial q_A}{\partial \underline{b}} = a - p,$$

which is negative since $a < p$, and the first partial derivative with respect to $\overline{b}$ is

$$\frac{\partial q_A}{\partial \overline{b}} = -\frac{(\delta-2)^2\overline{b} + a(\delta(\delta+2)-4) - 6\delta p + 8p}{6\delta - 8},$$

which is sure to be positive since $\delta < 1$. For the game in which B moves first, it initiates when $b < b_0$, where

$$b_0 = \frac{\delta(3\delta-4)\overline{b} + a(\delta(\delta+2)-4) - 2(3\delta-4)(p-q)}{2\left(\delta^2-2\right)},$$

such that the probability of initiation is

$$\Pr(\text{Initiate}) = \frac{b_0 - \underline{b}}{\overline{b} - \underline{b}}.$$

The first partial derivative with respect to $\overline{b}$ is

$$\frac{\partial \Pr(\text{Initiate})}{\partial \overline{b}} = -\frac{(\delta-2)^2\underline{b} + a(\delta(\delta+2)-4) - 2(3\delta-4)(p-q)}{2\left(\delta^2-2\right)\left(\overline{b}-\underline{b}\right)^2},$$

which is negative when

$$q > \frac{-a(\delta(\delta+2)-4) + 6\delta p - 8p}{6\delta-8} \quad \text{and} \quad \underline{b} < \frac{2(3\delta-4)(p-q) - a(\delta(\delta+2)-4)}{(\delta-2)^2}$$

and weakly positive if either constraint is not met. Finally, the first partial derivative with respect to $\underline{b}$ is

$$\frac{\partial \Pr(\text{Initiate})}{\partial \underline{b}} = \frac{(\delta-2)^2\overline{b} + a(\delta(\delta+2)-4) - 2(3\delta-4)(p-q)}{2\left(\delta^2-2\right)\left(\overline{b}-\underline{b}\right)^2},$$

which is positive when

$$q > \frac{-a(\delta(\delta+2)-4) + 6\delta p - 8p}{6\delta-8} \quad \text{and} \quad \overline{b} < \frac{2(3\delta-4)(p-q) - a(\delta(\delta+2)-4)}{(\delta-2)^2},$$

and weakly negative if either constraint is not met. □

Proposition 4 shows that both players' incentives to initiate crises depend not on the fact *that* information has been revealed (the important factor in Proposition 3) but on *what* information has been revealed. A is less willing to initiate as $\underline{b}$ increases, i.e. as B enters a crisis with a less resolute reputation, but moret willing to initiate as $\underline{b}$ decreases. Likewise, the probability with which B itself initiates the crisis can't be characterized solely in terms of his reputation; the effect of increases in both $\underline{b}$ and $\overline{b}$ can be positive or negative, depending on values of the status quo and the opposite bound of the type distribution. Thus, while our theoretical model leads us to expect that the probability that crises escalate should decline over time, it can generate *any* relationship—positive, negative, or even null—between the passage of time and the onset of crises, depending on the specifics of A's beliefs and B's share of the status quo.

Empirical Analysis

This table reproduces the key models from Table 4 of our main analysis with the LEAD data (Ellis, Horowitz and Stam 2015).

Table 1: Estimating Dispute Escalation within Different Samples (LEAD)

	Model 9		Model 10		Model 11		Model 12	
	Rivals		Contiguous		Politically relevant		All	
Dispute escalation								
Days since leader change	–0.222***	(0.078)	–0.159**	(0.078)	–0.072	(0.072)	–0.049	(0.066)
Days since regime change	–0.062	(0.041)	–0.056*	(0.033)	–0.095***	(0.031)	–0.094***	(0.028)
Joint democracy	–0.410	(0.345)	–0.440	(0.298)	–0.308	(0.272)	–0.460*	(0.243)
Democracy	–0.309*	(0.163)	–0.095	(0.161)	–0.251*	(0.133)	–0.161	(0.118)
Job insecurity	1.541**	(0.770)	–0.031	(0.803)	0.084	(0.747)	0.320	(0.668)
Age	0.007	(0.004)	0.006	(0.004)	0.003	(0.004)	0.004	(0.003)
Previous times in office	–0.214*	(0.123)	–0.139	(0.096)	–0.102	(0.085)	–0.085	(0.082)
Constant	2.860***	(0.805)	2.639***	(0.613)	2.084***	(0.637)	1.614***	(0.569)
Dispute onset								
Days since leader change	0.208***	(0.036)	0.148***	(0.024)	0.102***	(0.019)	0.081***	(0.016)
Days since regime change	–0.009	(0.015)	–0.002	(0.012)	–0.003	(0.013)	0.043***	(0.013)
Joint democracy	0.319*	(0.183)	–0.044	(0.139)	–0.066	(0.115)	–0.028	(0.084)
Democracy	0.050	(0.086)	–0.111	(0.076)	–0.173***	(0.065)	–0.194***	(0.052)
Job insecurity	–0.182	(0.513)	0.613*	(0.336)	0.730***	(0.244)	1.144***	(0.193)
Age	–0.001	(0.002)	0.001	(0.002)	0.005***	(0.002)	0.007***	(0.001)
Previous times in office	–0.011	(0.046)	0.038	(0.034)	0.002	(0.031)	–0.007	(0.027)
Other rivalries	–0.040*	(0.022)	0.005	(0.017)	0.005	(0.015)	0.080***	(0.012)
Distance	–0.013	(0.011)	–0.024***	(0.009)	–0.087***	(0.007)	–0.155***	(0.005)
Peace duration	–0.120***	(0.020)	–0.137***	(0.012)	–0.042***	(0.003)	–0.087***	(0.006)
Peace duration^2	0.004***	(0.001)	0.004***	(0.001)	0.000***	(0.000)	0.002***	(0.000)
Peace duration^3	–0.000*	(0.000)	–0.000***	(0.000)	–0.000***	(0.00)	–0.000***	(0.000)
Constant	–2.174***	(0.257)	–2.072***	(0.172)	–2.176***	(0.141)	–2.582***	(0.123)
ρ	–0.478***		–0.434***		–0.185*		–0.012	
Model χ^2	22.24		23.68		28.24		29.45	
Selection stage N	6199		19932		64255		550237	
Outcome stage N	629		867		1083		1316	

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

References

Ellis, Cali Mortenson, Michael C Horowitz and Allan C Stam. 2015. "Introducing the LEAD Data Set." *International Interactions* 41(4):718–741.

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

Powell, Robert. 1999. *In the Shadow of Power*. Princeton: Princeton University Press.

Wolford, Scott. 2007. "The Turnover Trap: New Leaders, Reputation, and International Conflict." *American Journal of Political Science* 51(4):772–788.