The Arbiter's Gambit

Jorge D. Ramos-Mercado*

March 28, 2025

Abstract

Pacifying interventions like ceasefires, strike bans, and mandated mediation aim to de-escalate conflict. But by lowering the immediate cost of disagreement, they can dull incentives to settle—paradoxically prolonging disputes. This paper formalizes that mechanism in a dynamic reputational bargaining model with indivisible stakes and evolving leverage. Unlike canonical models, where delay vanishes with perfect information and zero frictions, delay here persists due to pacifying policies themselves. Even in frictionless settings, interventions distort concession dynamics and induce inefficient delay. The framework applies broadly—from war and litigation to labor disputes—and yields a sharp empirical prediction: interventions that reduce conflict intensity increase dispute duration. We test this using a 200-year panel of 92 interstate wars and an instrumental variables strategy exploiting exogenous variation in military preparedness and domestic unrest. The results confirm the theory: pacifying interventions systematically prolong conflict, even as they reduce short-run costs.

1 Introduction

Why do conflicts persist—even when power asymmetries are large, information is complete, and institutional mechanisms are in place? Standard bargaining theory offers a clear prediction: stronger actors impose terms, weaker actors concede, and institutional intervention accelerates agreement. In this view, mechanisms like ceasefires, strike bans, and mandated mediation reduce the marginal cost of confrontation, thereby making compromise more attractive and delay less likely.

Yet in practice, such interventions frequently fail to resolve disputes. Wars, labor negotiations, and legal conflicts often drag on despite the presence of formal mediators. Weaker actors routinely extract significant concessions long after initial power dynamics appeared decisive. These patterns are at odds with canonical models, which predict that asymmetric power and institutional support should facilitate, not delay, resolution.

*EGADE Business School, Tecnológico de Monterrey. Email: jdramosmercado@tec.mx.

This paper departs sharply from standard theory. In canonical models—such as Fearon (1995), Powell (2004), and Abreu and Gul (2000)—inefficient delay arises from private information, belief noise, or commitment problems, and disappears as these frictions vanish. By contrast, the model generates delay even under perfect information and vanishing frictions. The inefficiency stems not from misunderstanding or noise, but from institutional efforts to pacify conflict. These interventions dull concession incentives and distort the strategic logic of bargaining, even in fully transparent settings.

This paper identifies a mechanism that explains this paradox. By reducing the short-run costs of disagreement, pacifying interventions lower the urgency to settle, thereby distorting concession incentives and inducing strategic delay—even in the absence of classical frictions.

We formalize this mechanism in a dynamic reputational bargaining model between two players negotiating over an indivisible surplus as conditions evolve. Unlike models where delay stems from hidden types or belief frictions, we show that inefficient delay persists even with perfect information and no noise—driven solely by pacifying policies that dull the urgency to settle. The framework captures three core elements of high-stakes institutional bargaining: reputational intransigence (whether genuine or strategic), indivisible stakes (such as sovereignty or corporate control), and shifting leverage over time.

The logic generalizes beyond war. In labor disputes, for instance, legal strike bans during arbitration reduce the short-run cost of disagreement. This makes both firms and unions more willing to hold out, hoping that future shifts in sentiment or internal dynamics improve their position. Even without uncertainty, delay emerges from the institutional softening of confrontation.

The logic generalizes to legal contexts. In complex litigation, court-imposed mediation or stay orders often delay trial proceedings to facilitate settlement. But when the immediate costs of litigation are deferred, actors may strategically prolong negotiations—anticipating that judicial turnover or evolving precedent might improve their position. The Apple v. Samsung patent case illustrates this dynamic: despite early rulings, the dispute persisted for seven years, with each side leveraging procedural slowdowns to extract concessions.

A central implication is that interventions aimed at reducing conflict intensity can systematically increase its duration. By lowering the cost of delay, such policies dampen incentives to concede. Like interest-free credit in a negotiation, they make waiting more attractive. This inverts a core prediction of canonical models: even as reputational frictions vanish, delay persists so long as pacifying interventions remain in place.

This result uncovers a dynamic inconsistency in institutional design. Interventions that reduce the risk of escalation may appear optimal ex post, under immediate pressure. But ex ante, they mis-align strategic incentives—weakening concession dynamics and encouraging actors to initiate or sustain disputes they would otherwise avoid.

The mechanism aligns with broader commitment failures in policy design. Financial bailouts and antitrust forbearance stabilize crises in the short term but encourage risky or antic-ompetitive behavior ex ante. Similarly, once peacekeeping or mediation policies are enacted, institutional and political frictions often make them difficult to unwind—regardless of their long-run effectiveness.

We test the model's predictions using a 200-year panel of 92 interstate wars. A panel IV strategy identifies the causal effect of pacifying interventions on conflict duration by leveraging slow-moving variation in domestic instability and baseline military preparedness. Both instruments pass weak-IV diagnostics and are plausibly orthogonal to wartime shocks.

The empirical findings confirm the theory: negotiations and pauses in fighting, while individually associated with reduced combat intensity, jointly extend the total duration of conflict. This interaction effect—central to the model—persists across historical periods, passes placebo tests in non-war disputes, and remains robust to excluding global hegemons.

Because interventions are often endogenously triggered by conflict dynamics, identification is critical. The IV approach isolates plausibly exogenous variation and shows that de-escalation efforts, far from resolving disputes, can entrench them.

Main Contribution. This paper make contributions to three areas of study. First, it identifies a novel commitment problem—affecting institutional mediators rather than disputants. Second, it develops a general framework for strategic delay in institutionalized bargaining under indivisibility and evolving power. Third, it provides new empirical evidence, across two centuries, that well-meaning interventions can backfire—prolonging rather than resolving conflict.

The remainder of the paper is organized as follows. Section 2 situates the paper within existing work on bargaining theory, institutional design, and conflict persistence. Section 3 introduces the formal model, while Section 4 characterizes the equilibrium and derives the main comparative statics. Section 5 presents the empirical analysis, testing the model's core predictions using a 200-year panel of interstate wars, while a placebo test to assess robustness is delegated to the appendix. Section 6 concludes.

2 Related work

Why do strategic actors engage in costly disputes rather than settling immediately? This foundational question spans political economy, international relations, and law and economics. The classical view—traced to Schelling (1967) and formalized by Fearon (1995)—portrays conflict as a costly signaling game driven by incomplete information. Wars, strikes, and lawsuits arise when actors misjudge an opponent's strength or resolve. Yet these models struggle to explain prolonged disputes in settings where capabilities are transparent—such as asymmetric wars, protracted litigation, or drawn-out labor standoffs.

A second tradition emphasizes limited commitment. Without enforceable agreements, actors may initiate conflict to avoid future losses. Powell (2004) shows that declining powers may preemptively strike to lock in favorable terms. Building on this, models by Fearon (2013), Filson and Werner (2002), and Leventoglu and Stanchev (2007) introduce reputational screening, where strong types separate from weak ones through costly delay. These approaches typically assume an institutional vacuum, where the absence of third-party enforcement sustains inefficiency.

Yet many real-world disputes occur under institutional oversight—courts, arbitrators, and peacekeepers regularly intervene. This raises a puzzle: why do conflicts persist even with credible mediators? Work by Beardsley (2011) and Moffitt (2013) suggests one answer: interventions can reduce the immediate cost of disagreement, enabling strategic delay. Relatedly, Tirole (2015) shows how institutions intended to stabilize outcomes may inadvertently generate strategic slack, while Alesina and Tabellini (2007) illustrate how long-term policy frictions can emerge from short-run political constraints.

Parallel literatures examine the role of indivisibilities and reputation. When the surplus cannot be divided, as in sovereignty claims or executive control, bargaining may fail altogether (Jackson and Morelli, 2009). Others highlight reputational incentives: some actors are obstinate with positive probability, and rational opponents may mimic this intransigence to extract better terms (Abreu and Gul, 2000; Fanning, 2021; Ekmekci and Zhang, 2024). However, most such models predict that inefficiencies vanish as frictions—like belief noise or communication delays—disappear.

This paper departs from that result. It shows that inefficiencies can persist even in frictionless settings when indivisibilities and institutional interventions interact. Pacifying policies, by lowering the cost of delay, dull incentives to settle. This generates inefficient delay not from misperception or commitment problems, but from rational strategic behavior in a softened bargaining environment.

The contribution is twofold. First, it extends reputational bargaining theory by identifying a new mechanism through which delay survives without frictions—so long as institutional interventions and indivisible stakes are present. Second, it yields a clean empirical prediction: inefficiency arises *only* under pacifying policies. This distinguishes the model from canonical delay theories and unifies prior insights across law, diplomacy, and labor relations into a tractable, testable framework.

3 Model Setup

This section presents a dynamic model of conflict where two sides are locked in a high-stakes negotiation over something that cannot be divided—like territory, executive control, or a legal ruling. The model centers on two forces that shape how disputes unfold. First, there's reputational uncertainty: some actors are truly inflexible due to ideology or politics, while others pretend to be in order to extract better terms. Second, institutional interventions—like ceasefires or mandated mediation—can reduce the short-term pain of conflict but unintentionally make delay more attractive. These dynamics are not specific to war; they also apply to litigation, strikes, or legislative standoffs.

The model proceeds in three stages. First, it builds a framework where players bargain over indivisible outcomes while conditions—such as the cost of fighting or the likelihood of victory—evolve over time. Crucially, these conditions are periodically altered by institutional efforts to soften the conflict. Second, we show that players tend to hold out during stable periods and only make concessions during key turning points—like when reputations collapse or power shifts. Third, we derive the model's core predictions. Specifically, we show that pacifying interventions actually lengthen disputes by dulling the incentive to settle, reduce the overall efficiency of outcomes, and—under certain conditions—make conflict unavoidable even when both sides fully understand each other's intentions.

3.1 Dynamic Effects of Fighting

We now formalize how conflict unfolds dynamically between two players, accounting for shifts in intensity and the opportunity for decisive outcomes.

Two players, 1 and 2, negotiate over a surplus normalized to one. Time is continuous and extends indefinitely. The dispute alternates between high and low conflict intensity, denoted by the exogenous process $\psi_t \in \{0, 1\}$, where high-intensity phases ($\psi_t = 1$) correspond to active confrontation (e.g., military combat, labor strikes, legislative gridlock), and low-intensity phases ($\psi_t = 0$) reflect de-escalated moments such as ceasefires, court mediation, or private negotiations. We interpret ψ_t as reflecting institutional interventions that dampen the immediate costs of disagreement.

Each player $i \in \{1, 2\}$ can, with intensity $\nu_{it} \geq 0$, secure a decisive victory. This may represent a military breakthrough, favorable court ruling, or legislative maneuver granting unilateral control. Decisive outcomes can only occur during high-intensity phases; hence, their arrival rate is $\psi_t \nu_{it}$. In addition, during high-intensity conflict, player i imposes flow costs $\psi_t c_{it}$ on their opponent. Both decisive outcomes and inflicted costs are thus conditional on conflict intensity.

The parameters $\{\nu_{it}, c_{it}\}_{i=1}^{2}$ evolve via a continuous-time, finite-state Markov process with $n < \infty$ states. Let $\ell = 1, \ldots, n$ be the current state. In state ℓ, player i has characteristics $(\nu_{i\ell}, c_{i\ell}) \in \mathbb{R}^2_+$. State transitions occur only forward (from ℓ to $\ell' > \ell$), at rate $\lambda_{t\ell\ell'}$ that depends on intensity: if $\psi_t = 1$ (high intensity), the transition rate is $\lambda_{\ell\ell'} \geq 0$, and if $\psi_t = 0$ (low intensity), it is $\hat{\lambda}_{\ell\ell'} \in [0, \lambda_{\ell\ell'}]$. We assume ψ_t is adapted to this process and remains constant between transitions.

This structure allows for dynamic, irreversible shifts in bargaining power. In war, prolonged fighting may erode a stronger party's advantage, enhancing the opponent's leverage. In labor disputes, sustained strikes can cause reputational harm, customer loss, and operational risk to firms, thereby improving union bargaining power. In each case, conflict reshapes the underlying incentives.

To simplify, we assume player 1 starts advantaged: $(\nu_{11}, c_{11}) \gg (\nu_{21}, c_{21})$, and that $\nu_{i\ell} < c_{j\ell}$ for all $i \neq j$ and ℓ, ensuring conflict is costly to both. Unlike models such as Krainin et al. (2020), which rely on private information about strength, our framework is suitable for disputes—like conventional war, litigation, and union strikes—where parties observe each other's capabilities. Here, delay emerges from strategic behavior under evolving conditions rather than hidden types.

3.2 Actions

At any moment, players may demand the full surplus ($\omega_{it} \in \{0, 1\}$), concede to their opponent, or irreversibly break off negotiations. Breaking off reverts the conflict to high intensity ($\psi_s = 1$ for all $s \geq t$), with real-world parallels including battlefield reengagement in war, court trial in litigation, and escalation to strike or lockout in labor disputes.

Demands are binary, reflecting the indivisibility of the contested good. This structure is appropriate for sovereignty claims like Jerusalem or Taiwan, corporate or legislative control, or union demands for pension transitions. Such indivisibilities constrain compromise, making mediation less effective. Arbitrated settlements, as studied in Fanning (2021), are often ineffective when control cannot credibly be divided. Even peaceful negotiation incurs costs, such as legal fees, reputational risks, and organizational strain, which reinforce the model's central trade-offs.

3.3 Intransigence

Each player is independently drawn as *obstinate* with probability $\mu \in (0, 1)$ at time $t = 0$. Obstinate players always demand the full surplus, break off negotiations at a constant rate $\phi > 0$, and are unresponsive to incentives. This captures actors constrained by ideological rigidity, religious absolutism, or intransigent constituencies.

Because the surplus is indivisible, obstinacy is unambiguous: players make maximal demands and eventually withdraw. Strategic (non-obstinate) players may mimic this behavior to extract concessions, generating classic reputational delay—even under full information.

3.4 Payoffs

The cost of conflict for player i up to time t is given by:

$$C_{it} = -\int_0^t \left(\bar{c}\,\eta_s + \psi_s\, c_{js}\right) e^{-rs}\,\mathrm{d}s,$$

where η_s indicates whether communication is open, ψ_s denotes conflict intensity, and $r \geq 0$ is the common discount rate. The constant $\bar{c}$ captures baseline costs of negotiation failure and is chosen large enough to induce delay:

$$\bar{c} > \phi + \max_{\ell} \sum_{\ell'} \lambda_{\ell\ell'}, \quad \text{and} \quad \bar{c} > \max_{\ell}\{\nu_{1\ell}, \nu_{2\ell}\}.$$

Players are symmetric except for ν_{it} and c_{it}, which lets us isolate the effects of reputation and shifting leverage. The game ends either in decisive victory or settlement. If player i wins at time t, her payoff is $e^{-rt} - C_{it}$, while player j receives $-C_{jt}$. If the players settle at time t and assign share $\omega_{it} \in \{0, 1\}$ to player i, then payoffs are:

$$\text{Player } i: \quad e^{-rt}\omega_{it} - C_{it}, \quad \text{Player } j: \quad e^{-rt}(1-\omega_{it}) - C_{jt}.$$

Settlements are instantaneous and override potential breakdowns. If both players concede simultaneously, the surplus is destroyed.

3.5 Strategies and Equilibrium

Let h_0 denote the initial history. At time $t > 0$, the history h_t records the $m \in \{0, \dots, n\}$ state transitions observed up to time t, given by: $h_t = \big((t_1, \ell_1), \dots, (t_m, \ell_m)\big)$, with $t_1 < \cdots < t_m$, where each ℓ_k indicates the state reached at time t_k.

A strategy for player i is a stopping time τ_i with respect to the filtration $\{h_s\}_{s \leq t}$, denoting the moment when player i chooses to concede—conditional on communication remaining open, no settlement having been reached, and no decisive outcome occurring.

Let $\mu_{jt}(h_t)$ represent player i's belief at time t that opponent j is obstinate, conditional on history h_t. These beliefs evolve endogenously via Bayes' Rule, updating based on observed behavior.

We adopt the Perfect Bayesian Equilibrium (PBE) concept: strategies must be sequentially rational, and beliefs must be updated consistently.

Equilibrium Structure. We take this opportunity to present the equilibrium, but note that the formal derivations are presented in later sections and in the appendix. The result goes as follows.

Theorem 1. *In equilibrium, at most one player concedes at time $t = 0$. Otherwise, the dispute evolves as a war of attrition until one of the following occurs:*

i. Beliefs converge to 1,

ii. A decisive resolution occurs,

iii. Communication breaks down, or

iv. A state transition occurs.

If a state transition occurs, one player concedes immediately with positive probability. Otherwise, the war of attrition continues with updated concession dynamics. Each of these events ends the strategic phase.

3.6 Example 1: State Shifts and Strategic Concession

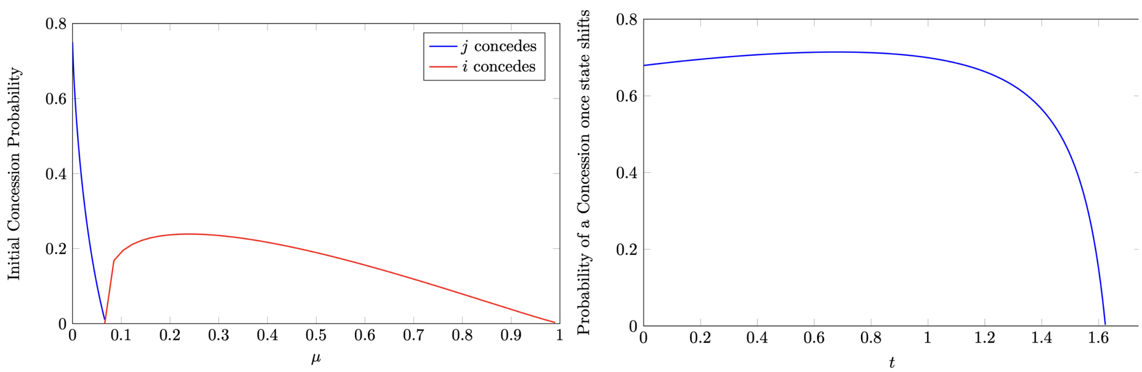
While the next section formally characterizes general model dynamics, the complexity may obscure key distinctions from canonical reputational bargaining. To address this, we present a stylized numerical example in which the equilibrium can be visualized clearly. This example assumes players are forced to negotiate in a low-intensity environment, thereby illustrating how policies that suppress conflict intensity can unintentionally backfire.

Assume $n = 2$, $\psi_t = 0\ \forall t$, $r = 0$, and $\lambda_{12} = \lambda = 1$. Conflict parameter values are

$$(c_{i1}, \nu_{i1}, c_{j1}, \nu_{j1}) = (1, 1/2, 1/2, 0), \quad (c_{i2}, \nu_{i2}, c_{j2}, \nu_{j2}) = (2/5, 0, 2, 1/5),$$

with baseline negotiation costs $\bar{c} = 2$, obstinacy hazard $\phi = 1$, and initial reputation for obstinacy $\mu \in [0, 1]$. Player i initially holds the upper hand, imposing higher costs on j and being more likely to win outright. However, if the state shifts, this advantage reverses.

Figure 1 illustrates two key dynamics. Panel (a) shows that when players' initial reputations for obstinacy are low, the initially weaker player j concedes with strictly positive probability—declining in μ. This is because j expects to be locked in a disadvantageous



(a) Initial concession probability (b) State shift concession probability

Figure 1: Discontinuous concession probability at time $t = 0$ (left panel) and following a state shift at unstable times $t > 0$. Panel (b) assumes low μ, such that j concedes from the outset.

position for a prolonged period. As $\mu \to 0$, the probability of early resolution converges to 0.75, sharply contrasting with canonical reputational models where this limit approaches 1. This divergence arises from the introduction of state dynamics and indivisible outcomes, which fundamentally alter early-move incentives.

Panel (b) plots the probability that j concedes after a state transition, assuming a low μ. The probability remains relatively stable until a critical time $T > 0$, at which beliefs converge to 1 and concession becomes inevitable.

This example clarifies the interaction between reputation and the threat of state shifts. When μ is high, player i (initially advantaged) is more likely to concede immediately, while j concedes gradually if no immediate resolution occurs. If a state shift does occur, j may concede probabilistically, triggering a renewed war-of-attrition where i concedes at a faster rate. For low μ, however, j lacks the credibility to compel early concessions from i, and delay persists.

3.7 Base Result 1: Payoffs Following Communication Breakdown

When negotiations fail, disputes are resolved externally—via military outcomes in war, judicial rulings in litigation, or institutional intervention in labor disputes. This subsection characterizes the payoffs strategic players receive once communication breaks down, locking both parties into high-intensity conflict until a decisive outcome is imposed.

Lemma 1. *If communication breaks down at time t, then $\psi_s = 1$ for all $s \geq t$, and player*

i's continuation payoff in state ℓ is $-B_{i\ell}$ where

$$B_{i\ell} = \frac{(c_{j\ell} - \nu_{i\ell}) + \sum_{\ell' \neq \ell} \lambda_{\ell\ell'} B_{i\ell'}}{r + \nu_{i\ell} + \nu_{j\ell} + \sum_{\ell' \neq \ell} \lambda_{\ell\ell'}}. \tag{1}$$

Given that $c_{j\ell} > \nu_{i\ell}$ for all i, j, ℓ, it follows that $B_{i\ell} > 0$.

Equation (1) represents the continuation payoff a player receives if communication breaks down and the dispute enters a phase of high-intensity conflict. This expression captures the discounted cost of being locked into an external resolution process: the player faces ongoing losses from her opponent's ability to inflict harm, partially offset by the probability of achieving a decisive victory or transitioning into a more favorable bargaining state. Because this continuation value is strictly negative, strategic players always prefer to settle or concede before communication breaks down. As a result, only truly obstinate players—those unwilling to compromise regardless of cost—ever allow bargaining to collapse. This insight has two key implications. First, it highlights the deterrent role of breakdown risk in shaping strategic behavior. Second, it explains why re-opening communication is never optimal in equilibrium: doing so would simply induce immediate concession by a rational player seeking to avoid renewed exposure to high-intensity conflict.

Equation (1) further illustrates how the prospect of power shifts exacerbates losses. In absorbing states (i.e., $\lambda_{\ell\ell'} = 0$ for all $\ell' \neq \ell$), $B_{i\ell}$ simplifies to the present value of costs net of victory probabilities. In transient states, however, future power transitions expose players to additional downside risk. A player may initially dominate, but if her strategic position is eroded by endogenous changes, her long-run payoff can deteriorate sharply.

Using the example provided in sub-section 3.6, the payoffs and their dynamic nature can be best illustrated. Suppose $n = 2$, $r = 0$, and $\lambda_{12} = \lambda > 0$. In state 1, player i is advantaged: $(c_{i1}, \nu_{i1}) = (1, 1/2)$ and $(c_{j1}, \nu_{j1}) = (1/2, 0)$. In state 2, the roles reverse: $(c_{i2}, \nu_{i2}) = (2/5, 0)$ and $(c_{j2}, \nu_{j2}) = (2, 1/5)$. Calculating payoffs yields $B_{j2} = 1$ and $B_{i2} = 10$, while

$$B_{i1} = \frac{20\lambda}{1 + 2\lambda}, \quad B_{j1} = \frac{4(1 + \lambda)}{1 + 2\lambda}.$$

If $\lambda > 1/4$, then $B_{i1} > B_{j1}$: despite initially imposing higher net costs, player i is worse off due to the risk of transitioning into a disadvantageous future state.

These dynamics apply across institutional settings. In war, a technologically superior army may lose its edge if a weaker opponent captures its weapons or exploits its vulnerabilities. In politics, a dominant coalition may suffer reputational costs as prolonged standoff erodes public support. In labor or corporate disputes, firms may lose leverage over time as sustained disruption weakens their market position.

The Apple vs. Samsung patent dispute exemplifies these dynamics. In 2011, Apple sued Samsung for over $2 billion, bypassing mediation and committing to litigation. Samsung responded with counterclaims and appealed all the way to the U.S. Supreme Court, which ultimately ruled in Samsung's favor. Although Apple won partial damages in lower courts, it secured only $539 million by 2018—far less than initially sought. This case underscores how strategic missteps and prolonged escalation, even by initially advantaged parties, can diminish returns from conflict and expose both sides to mounting risks.

4 Equilibrium Characterization

This section characterizes equilibrium behavior and derives key comparative statics. As in standard reputational bargaining models (e.g., Abreu and Gul, 2000; Fanning, 2021), at most one player concedes at the outset. If no immediate concession occurs, the dispute unfolds as a war of attrition. The innovation in this framework lies in the interaction between reputational dynamics and state transitions: at most one player concedes immediately following a shift in the state, and the anticipation of such transitions influences concession behavior during stable periods.

4.1 Base Result 2: Concession Dynamics

To make our characterization precise, we start by defining two key concepts. First, a time t is called *stable* if $t > 0$ and the state does not change over the interval $[0, t)$. That is, $h_t = \cap_{s\in[0,t)} h_s$. Otherwise, t is said to be *unstable*. Second, we define a time t as *strategic* if four conditions are met: (i) communication has not broken down, (ii) no agreement has been reached, (iii) neither player has achieved a decisive external victory, and (iv) both players still assign less than full certainty to their opponent being obstinate, meaning $\max_i \mu_{it} < 1$.

We now present the central equilibrium result:

Lemma 2. *Every equilibrium satisfies the following properties:*

(i) After each unstable, strategic time $t > 0$, there exists a duration $T_t > 0$ such that, if no further state transitions occur during $(t, t + T_t]$, both players' beliefs converge: $\mu_{it} \to 1$.

(ii) At most one player concedes immediately, whether at $t = 0$ or at any unstable, strategic time $t > 0$.

(iii) No player concedes abruptly at a stable, strategic time $t > 0$.

(iv) At stable, strategic times $t > 0$, players randomize between conceding and demanding the surplus.

Interpretation: Lemma 2 outlines how concessions unfold in equilibrium and how they differ from traditional reputational bargaining models. In classical setups, disputes end once beliefs converge. In contrast, our model introduces uncertainty through potential shifts in bargaining conditions. When such a shift (or *state transition*) occurs, it may prompt an immediate concession. If not, players enter a prolonged standoff—updating beliefs slowly about whether the other side will hold firm or give in.

The core insight is this: concession behavior is shaped not just by belief dynamics but also by anticipation of future changes. Players delay in hope that a new state might improve their bargaining position. Over time, belief convergence increases pressure to concede—but that pressure resets each time a state transition occurs. This produces a more realistic negotiation pattern: long stalemates interrupted by sudden concessions. In stable periods, players engage in a war-of-attrition. The pace of concession depends on the current state and expected costs of remaining in conflict, captured by $B_{i\ell}$ (see Lemma 1). Expecting a state change, players slow down their concessions, making the timing asymmetric even under similar reputational frictions.

As a result, even when both players face identical environments, one may concede earlier. Only in completely symmetric cases—where $\nu_{1\ell} = \nu_{2\ell}$ and $c_{1\ell} = c_{2\ell}$ for all ℓ—do these asymmetries vanish, leading to rare or degenerate concessions. The full proof of Lemma 2 is in the online appendix. To support our analysis, we introduce some notation. Let $C_{i\ell t}$ denote the rate at which player i concedes at a stable, strategic time $t > 0$ in state ℓ. Let $1 - q_{i\ell\ell' t}$ represent the probability that player i concedes immediately after an unstable time $t > 0$, following a transition from state ℓ to a more favorable state $\ell' > \ell$. For brevity, we write q_{it} when the specific state transition is clear from the context.

4.2 Dynamics and Concession Behavior

This section derives the core expressions that describe equilibrium behavior in the model. Specifically, we focus on four key objects: the continuous concession rates $C_{i\ell t}$, the discrete concession probabilities $q_{i\ell\ell' t}$, the belief trajectories $\mu_{i\ell t}$, and the convergence duration T_t. These expressions are characterized for each player i, state ℓ, and strategic time t. Together, they form the backbone of how strategic delay and concession decisions evolve over time under uncertainty and shifting environments.

4.3 Equilibrium Payoffs and Concession Dynamics

Let $W_{i\ell t}$ denote player i's expected discounted payoff at a strategic, stable time $t > 0$, conditional on being in state ℓ. This value summarizes the tradeoff a player faces between conceding immediately or continuing the dispute. In equilibrium, $W_{i\ell t}$ satisfies a recursive relationship that reflects several key factors: the ongoing costs of negotiation, uncertainty about the opponent's reputation, the probability of securing a decisive victory, and the likelihood of state transitions. A full derivation of this relationship can be found in Appendix A.

Under standard assumptions, we derive a concise expression for the *total concession rate*:

$$K_{i\ell t} = \bar{c} + \mu_{jt}\phi B_{i\ell} + \psi_t(c_{i\ell} - \nu_{j\ell}). \tag{2}$$

This equation captures the interplay of three central forces driving concession behavior. First, the term $\bar{c}$ represents the baseline cost of bargaining delays—this creates persistent pressure to settle, even when immediate conflict intensity is low. Second, reputational concerns play a major role: when a player believes their opponent is likely to be obstinate (represented by a high μ_{jt}), and when the breakdown of communication is likely (probability ϕ), the player becomes more inclined to concede early to avoid being stuck in an inefficient impasse. This effect is magnified by the potential continuation cost $B_{i\ell}$. Third, the expression $\psi_t(c_{i\ell} - \nu_{j\ell})$ reflects each player's leverage: it compares the cost they can impose on their opponent during continued conflict with the risk of suffering a decisive loss. If leverage is high, delay becomes more attractive; if low, concessions become more appealing.

These three components jointly determine concession behavior in a dynamic and strategic setting. Unlike standard screening models, where delay arises mainly from learning about hidden types, our model shows that players delay due to evolving reputations and shifting risks. As a result, the total concession rate is always positive, reputations become more entrenched over time, and prolonged disputes can persist even in the absence of private information.

Corollary 3. *In any equilibrium, the total concession rate $K_{i\ell t}$ increases with the baseline negotiation cost $\bar{c}$, the conflict intensity ψ_t, the breakdown probability ϕ, the cost a player imposes on their opponent $c_{i\ell}$, and the opponent's continuation cost $B_{j\ell}$. It decreases with the player's discount rate r and the opponent's likelihood of achieving a decisive victory $\nu_{j\ell}$. The effect of the state transition rate $\lambda_{\ell\ell'}$ for $\ell' > \ell$ is generally ambiguous.*

This corollary highlights the drivers behind a player's inclination to concede. Concessions become more likely as the burdens of negotiation increase, as conflict becomes more intense, and as the opponent stands to lose more if bargaining fails. Conversely, greater patience and a higher risk of losing decisively reduce the incentive to concede. Even when reputational frictions are modest, the fundamental costs of negotiation generate sustained pressure to settle. This underlines one of the model's central insights: strategic delays are not solely the result of uncertainty about types, but also emerge from the anticipation of future changes in bargaining power.

4.4 Belief Dynamics

This section examines how player j's belief about player i's type—denoted by μ_{it}—evolves over time. The intuition is that maintaining a posture of toughness is costly, and over time, this gradually convinces player j that player i may in fact be obstinate, especially if the bargaining environment remains unchanged. Learning also occurs at the moment a state shift happens, particularly if no concession is made. In such cases, players re-evaluate whether it is worth continuing the dispute, considering how the new conditions will shape future bargaining.

We first consider belief updating at unstable times, when a state transition occurs. Suppose $t = 0$ or $t > 0$ is a *strategic, unstable* moment—meaning a state has just changed. Let μ_{it^-} denote player j's belief just before this shift. If player i concedes with probability $1 - q_{it}$, then observing a concession causes player j to update her belief to zero, since only non-obstinate types concede. If i does not concede, player j observes an action taken with probability q_{it} and applies Bayes' Rule to revise her belief as $\mu_{it} = \frac{\mu_{it^-}}{q_{it}}$. The value of q_{it} is determined endogenously in equilibrium and depends on the expected horizon T_t over which beliefs would otherwise converge to one if no additional state transitions occurred.

Now consider what happens at a *strategic, stable* time $t > 0$, where the state remains fixed at ℓ. Let μ_{it} again denote player j's belief about whether i is obstinate. Suppose no concession, breakdown, or state transition occurs during the interval $[t, t + \Delta_t]$. In this setting, an obstinate player will never concede but may break off communication at rate ϕ, whereas a strategic type will never break off communication but may concede at rate $C_{i\ell t}/(1 - \mu_{it})$. Applying Bayes' Rule, we can express the updated belief at time $t + \Delta_t$ as:

$$\mu_{it+\Delta_t} = \frac{\mu_{it}(1 - \phi\Delta_t)}{1 - \phi\mu_{it}\Delta_t - C_{i\ell t}\Delta_t} + o(\Delta_t),$$

where the term $o(\Delta_t)$ is of higher order and satisfies $\lim_{\Delta_t \searrow 0} o(\Delta_t)/\Delta_t = 0$. Taking the limit as $\Delta_t \to 0$, we obtain a differential equation that describes belief evolution continuously over time:

$$\frac{\mathrm{d}}{\mathrm{dt}} \ln \mu_{it} = C_{it} - \phi(1 - \mu_{it}). \tag{3}$$

Equation (3) implies that beliefs about player i's obstinacy increase over time whenever she does not concede. The term C_{it} captures the expected rate of concession for a strategic type, while the expression $\phi(1 - \mu_{it})$ accounts for the absence of communication breakdown—behavior more typical of an obstinate type. In equilibrium, this mechanism ensures that beliefs converge toward one with strictly positive probability before the next key event (such as a concession or a shift) occurs.

4.5 Discrete Concessions and Duration

This section characterizes the final moving parts of equilibrium: the probability of a discrete concession at a moment of instability, and the amount of time needed for reputational beliefs to converge to certainty afterward. Let $1 - q_{it}$ denote the probability that player i makes a discrete concession at a strategically unstable moment t, and let T_t denote the belief convergence time that follows. These objects are determined recursively via backward induction, and their formal derivation is deferred to Appendix A.2.

Equilibrium behavior during transitions is governed by a form of selection: Lemma 2 implies that at most one player concedes discontinuously at any given state shift. If $q_{it} < 1$, then $q_{jt} = 1$, and the burden of concession falls entirely on player i. This determination unfolds in three steps: establishing the convergence time T_t, identifying the player most likely to concede, and then computing the actual concession probability.

Although the mathematical characterization of this process involves a coupled system of beliefs, flows, and integral equations, the logic is intuitive. When a state change occurs, each player reevaluates the speed at which their opponent's reputation might stabilize. The player whose opponent becomes convinced of their obstinacy more quickly enjoys strategic leverage; the slower party faces growing pressure to concede. If both players begin with identical reputations—say, at time $t = 0$—the equilibrium favors the player who is expected to concede more rapidly over time. However, if a state change occurs closer to the end of the convergence period, reputations may have evolved asymmetrically, shifting the identity of the conceding player.

The concession probability itself reflects both the initial reputation imbalance and the difference in anticipated concession paths. If player i fails to concede at a moment when they are expected to, their opponent's belief adjusts discontinuously downward via Bayes' rule. The rate at which this belief recovers over time depends on both players' concession dynamics and the rate at which the dispute could collapse due to communication breakdown. These dynamics make it possible for the burden of concession to switch over the course of the game: the player who concedes first may not always concede again. Ultimately, the resolution of strategic uncertainty rests on the interaction of belief convergence rates and the timing of state transitions. The exact form of this interdependence is given in the recursive equilibrium solution, presented in the appendix A. In particular, to complete the characterization one must derive quantities using backwards induction. The precise recursive solution for these quantities is provided in Appendix. The derivation follows standard arguments via backward induction and is omitted here for brevity.

4.6 The Effect of Pacifying Policies on Dispute Persistence

The term ψ_t plays a central role in the model, as it captures exogenous interventions designed to limit the intensity of a dispute. Many real-world third-party interventions—such as ceasefires during war, court-mandated mediation in legal conflicts, or labor contracts that prohibit strikes—can be represented within the model as periods where $\psi_t = 0$ over some interval $[s, t]$ for $0 \leq s < t$, thereby reducing or eliminating direct high-intensity engagement between the disputing parties.

The primary motivation for such interventions is to reduce the expected duration of high-intensity conflict. However, an important question arises: do these reductions in expected intensity inadvertently produce longer or more persistent disputes? Notably, even as key frictions in the model vanish—specifically, as $(\mu, \bar{c}, \phi, \sup_{\ell,\ell',t} \lambda_{\ell\ell' t}) \to 0$—the probability of immediate resolution remains strictly less than one. Moreover, the expected duration of the dispute continues to decrease as ψ_t increases, suggesting a complex and non-linear relationship between intensity and resolution.

To formalize this idea, define $\Psi_t = \int_0^t \psi_{s|\ell=1}\,\mathrm{d}s$ as the cumulative time spent in high-intensity conflict, conditional on the state remaining at its initial value up to time t. A policy $\{\psi_t\}$ is said to be *pacifying* if it ensures that this cumulative exposure remains finite over time; that is, $\lim_{t\to\infty} \Psi_t < \infty$. The central result in this section shows that even under vanishing frictions, such pacifying policies can lead to persistent disputes.

To quantify this, define the limiting expected, discounted duration of the dispute under a given policy $\{\psi_t\}$ as $d(\{\psi_t\}) \equiv \lim_{(\mu,\phi,\sup_{\ell,\ell',t}\lambda_{\ell\ell' t})\to 0} \mathbb{E}_0[e^{-r\tau}]$, where τ denotes the stopping time at which the dispute ends. Let $q(\{\psi_t\})$ denote the limiting probability that the dispute does not resolve immediately at time $t = 0$. We now state the following lemma.

Lemma 4. *Fix some policy $\{\psi_t\}$. Then, as $(\mu, \bar{c}, \phi, \sup_{\ell,\ell',t} \lambda_{\ell\ell' t}) \to 0$, it follows that*

$$d(\{\psi_t\}) = 1 - r \int_0^\infty e^{-rt - \sum_i c_{i1}\Psi_s}\,\mathrm{d}t < 1 - e^{-\sum_i c_{i1} \lim_{t\to\infty} \Psi_t} \tag{4}$$

and

$$q_0(\{\psi_t\}) = e^{-[(\nu_{11}-\nu_{21})+(c_{11}-c_{21})]\lim_{t\to\infty}\Psi_t}. \tag{5}$$

This result reveals that even in the absence of standard sources of delay—such as reputational frictions, shifting power dynamics, or private information—pacifying interventions can still prolong disputes. By lowering the short-run costs associated with disagreement, such policies reduce the urgency to reach an agreement. Consequently, the probability of an immediate settlement falls, and the expected duration of the dispute increases. This introduces a crucial insight: well-meaning efforts to de-escalate conflict, such as ceasefires or mediation mandates, can backfire if they fail to account for how reducing conflict intensity reshapes the strategic calculus of the parties involved. The model thus uncovers a fundamental tension between short-term stabilization and the long-term resolution of disputes.

4.7 Bargaining Power and the Effect of Pacifying Policies

This section derives a closed-form expression for each player's bargaining power and examines how dispute intensity—captured by the policy sequence $\{\psi_t\}$—affects the distribution of bargaining outcomes. Because the surplus is indivisible, a player's bargaining power is defined as the probability that they ultimately secure the surplus as all frictions vanish. A central question is whether an initially disadvantaged player is inherently less likely to prevail. The analysis also considers how pacifying policies, which reduce either the intensity or duration of the conflict, influence the final allocation of bargaining power.

The findings reveal two important effects. First, conditional on the dispute not being resolved immediately, the initially disadvantaged player actually holds greater bargaining power. Second, if the intensity policy is pacifying—meaning it limits the cumulative time spent in high-conflict states—then the total bargaining power across both players is strictly less than one. This implies an inefficient equilibrium outcome, where part of the surplus is effectively lost due to unresolved bargaining.

To formalize these ideas, define the function $p_i(\mu, \bar{c}, \phi, \sup_{\ell,\ell',t} \lambda_{\ell\ell' t}, \ell, t)$ as the probability that player i ultimately secures the surplus—either by winning outright or because the opponent concedes—at any time $t > 0$ in state ℓ. That is,

$$p_i(\mu, \bar{c}, \phi, \sup_{\ell,\ell',t} \lambda_{\ell\ell' t}, \ell, t) = \mathbb{E}[\chi(i \text{ wins or } j \text{ concedes})]. \tag{6}$$

Then, player i's bargaining power is defined in the limit as frictions vanish:

$$p_i \equiv \lim_{(\mu,\phi,\sup_{\ell,\ell',t} \lambda_{\ell\ell' t}) \searrow 0} p_i(\ell, 0^+). \tag{7}$$

If the sum of both players' bargaining power is less than one, the outcome is considered wasteful, as part of the surplus fails to be allocated.

Lemma 5. *For any admissible conflict intensity policy $\{\psi_t\}$, the bargaining power of player*

$i \in \{1, 2\}$ is given by

$$p_i = \left(\frac{c_{i1}}{c_{i1} + c_{j1}}\right)\left(1 - e^{-\sum_i \nu_{i1} \lim_{T\to\infty} \Psi_T}\right). \tag{8}$$

Moreover, if $\{\psi_t\}$ is pacifying, then the outcome is wasteful, in the sense that $\sum_i p_i < 1$.

This result highlights that a player's bargaining power is not simply determined by their initial strength or position but rather by their capacity to impose sustained costs on their opponent. Players who can prolong disputes and make them more painful for the other side hold greater strategic leverage. Importantly, the final allocation of bargaining power also hinges on the nature of the conflict itself. When pacifying policies reduce the time spent in high-intensity conflict, they unintentionally reduce both players' bargaining power. The result is an inefficient outcome, where some surplus is effectively lost. Even when standard frictions disappear, such interventions create room for strategic delay. This insight underscores a broader theme of the model: efforts to de-escalate conflict in the short run can unintentionally undermine long-term resolution.

4.8 Example 2: Pacifying Policies can make dispute unavoidable

The previous example demonstrates that as the probability that players are non-strategic approaches zero, the probability of an initial concession does not converge to one. This raises two questions: is this result an artifact of the chosen numerical example?; does the fact power shifts give way to this example? To address these questions, we present a secondary example in which the state (for certain) does not shift. On one hand, this example cannot illustrate how concession behavior is affected by the expectation of a power shift. On the other hand, the example *does* clarify that as the reputational frictions vanish (i.e., as $(\mu, \phi) \to 0$), the probability that the conflict is resolved from the outset can converge to 0.

The example proceeds as follows. Assume that state 1 is absorbing, meaning that for all $\ell = 2, \ldots, n$, we have $\max_{t\geq 0} \lambda_{1\ell t} = 0$. In addition, suppose that for some player i, we have $B_{i1} > B_{j1}$, that $\psi_t = 0$ for all $t \geq 0$, and that $(\bar{c}, \phi) \gg 0$.

Lemma 6. *Suppose that $\lambda_{1\ell t} = \psi_t = 0$ for all $t \geq 0$, that $B_{i1} > B_{j1}$, $\mu < 1$, and that $(\bar{c}, \phi, \mu) \gg 0$. Then as reputational frictions go to zero (i.e., as $(\phi, \mu) \to 0$), the probability of an initial concession converges to* 0.

5 Empirical Section

This section empirically tests the model's central prediction: that institutional interventions—such as ceasefires, mandated mediation, and peace negotiations—can inadvertently prolong conflict by lowering the short-run costs of continued disagreement. While the theoretical framework applies broadly across domains of strategic bargaining, the empirical analysis focuses on interstate wars, where both intervention and conflict dynamics are observed with high fidelity over long historical horizons.

The empirical analysis proceeds in two parts. First, it assesses whether third-party interventions are causally associated with longer durations of active combat in a dataset of 92 interstate wars spanning the past 200 years. Second, it broadens the analysis to include militarized interstate disputes (MIDs) to examine whether interventions alter escalation probabilities in less formalized conflict settings.

To isolate plausibly exogenous variation in institutional involvement, the paper employs an instrumental variables (IV) strategy. Instruments are based on pre-determined proxies for domestic instability (long-run social unrest) and latent military capacity, capturing shifts in internal constraints and bargaining power that influence the likelihood of intervention—but are unlikely to be endogenous to short-term combat dynamics. The key outcome variable—duration of active combat—is chosen to closely track the theoretical construct of costly delay.

Because traditional survival models perform poorly with small sample sizes and low baseline hazards, I supplement the core analysis with panel regressions and robustness checks. The results are consistent with the model's predictions: institutional efforts to pacify conflict can, under certain conditions, delay resolution and entrench bargaining frictions. The remainder of this section outlines the data, empirical framework, main results, and their implications for the theory.

5.1 Data Description

This section presents the primary panel dataset used in the empirical analysis and outlines key trends relevant to the theoretical model. The foundation of the panel is a dataset of 92 interstate wars spanning 1823 to 2003, originally compiled by Min (2020), which includes day-level records of formal negotiations between belligerents. To this, we append detailed data on the timing of major battles from Min (2021), which identify when battles began and measure the duration of pauses between them. Finally, we integrate time-varying indicators of each side's war-sustaining capacity using national resource data from Lyall (2020), but the data and methodology to derive the instruments uses additional data and is described in subsequent sections.

These components correspond closely to the core constructs of the theoretical model. Periods of negotiation map to strategic episodes in which combatants exchange proposals aimed at peaceful resolution. In parallel, the model's binary function ψ_t captures the presence or absence of active combat, where $\psi_t = 1$ denotes fighting and $\psi_t = 0$ indicates a lull or pause. Accordingly, the data define $\psi_t = 0$ for the intervals between major battles and $\psi_t = 1$ for all other days. This coding implies that fighting may have occurred at the times when combatants fought, but these fights are only skirmishes and did not escalate significantly.

The negotiation variable is derived from a daily-level dataset that systematically codes all instances of direct or mediated wartime communication with the ostensible aim of reaching a settlement Min (2020). The battle-level data provide temporal precision, while the national resource indicators reflect both the evolving costs imposed by each side ($c_{i\ell}$ for $i = 1, 2$ and $\ell = 1, \ldots, n$) and the intensity with which decisive engagements occur. Importantly, although decisive battles in the real world may not immediately terminate a war, the model abstracts from this nuance—an assumption the empirical analysis is designed to accommodate.

Finally, this section focuses exclusively on interstate disputes rather than civil wars. Although civil wars are more frequent, they introduce an additional layer of complexity absent from the model: the risk that a non-governmental faction may refuse to honor a peace agreement. In contrast, sovereign states that renege on agreements face tangible international sanctions and diplomatic marginalization. In addition, intra-state actors are often fragmented, making it unclear who should be sanctioned if an agreement is violated and how such sanctions could be enforced. Given these challenges, interstate disputes provide a more direct empirical setting in which to test the model's predictions.

5.2 Summary Statistics Over Time

This section examines key summary statistics comparing the Early (1823–1914) and Modern (1914–2003) periods, as presented in Table 1. The table provides descriptive statistics on 92 interstate wars, highlighting significant shifts in the duration, negotiation patterns, and outcomes of conflicts over time. Wars have grown longer, with the average duration increasing from 331 to 456 days, while the time spent actively fighting has risen from 186 to 327 days. Despite this, the number of negotiations per war has doubled, and the time allocated to negotiations has more than tripled, from 34 to 113 days. However, the effectiveness of negotiations in ending wars has declined, with the share of wars concluding after successful talks dropping from 51% to 32%. Conversely, the share of wars ending after decisive battles has risen from 15% to 24%, suggesting a greater reliance on military outcomes in modern

	1823-1913	1914-2003
Number of Interstate Wars	41	51
Duration (days)		
Mean	331	456
Median	163	130
Time Spent Fighting (days)		
Mean	186	327
Median	104	117
Number of Negotiations		
Mean	2	4
Median	1	2
Time Spent in Negotiations (days)		
Mean	34	113
Median	6	18
Share of Disproportionate Wars (%)	40	33
War Outcomes (%)		
Ending after Effective Negotiation	51	32
Ending after a Decisive Battle	15	24
Other Outcome	34	44

Notes: A war is classified as *disproportionate* if one side has a CINC index at least 5 times greater than its opponent. A *negotiation* is considered effective if the war ends within a week of the negotiation's conclusion and a decisive battle did not also precede the outcome. A *battle* is classified as decisive if it ends with a clear winner less than a week before the war's official conclusion.

Table 1: Summary statistics on war outcomes, comparing the Early (1823–1914) and Modern (1914–2003) periods.

conflicts. In addition, the proportion of disproportionate wars—where one side possesses overwhelming power—has decreased, reflecting a more balanced distribution of military capabilities over time. These trends underscore fundamental changes in the nature of interstate wars, particularly in how they are fought and resolved.

A key distinction between the two periods lies in the emergence of intergovernmental organizations and their role in promoting ceasefires during peace negotiations. According to UN (2022), modern mediation efforts explicitly advocate for ceasefires as a critical component of the peace process. The institutionalization of ceasefires dates back to the Hague Convention of 1907, which provided the first internationally recognized definitions and guidelines for their use. Their effectiveness was further demonstrated at the end of World War I, reinforcing their role in conflict resolution (Davion, 2020). However, while ceasefires are often framed as de-escalation tools, the model suggests they may inadvertently prolong disputes by reducing the short-run costs of continued conflict—thereby weakening incentives to settle.

This paper hypothesizes, consistent with the theoretical model, that the increasing use of ceasefires has contributed to the declining efficacy of peace negotiations and the prolonged duration of wars. As ceasefires allow combatants to regroup and reassess their strategies, they may inadvertently reduce incentives for immediate conflict resolution, leading to prolonged disputes rather than swift negotiated settlements.

5.3 Overall Trends in Interstate Disputes

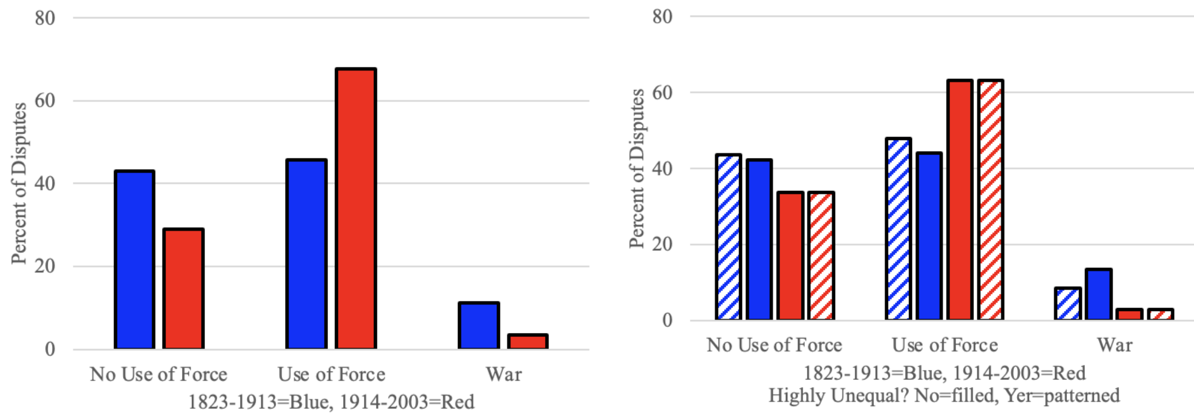
The summary statistics above indicate that in both periods, over 33 percent of interstate wars involved combatants with highly unequal military capacities. According to the model, if the balance of power remains stable, the stronger side is expected to achieve a decisive victory. If it fails to do so, the weaker opponent is more likely to extract concessions. More generally, the model predicts that, in the absence of power shifts, weaker combatants should concede without engaging in conflict. This implies that disputes between highly unequal opponents should escalate to war less frequently than disputes between more evenly matched states when power shifts are sufficiently rare.

To test this prediction, we use data from Sarkees and Wynman (2010) on militarized interstate disputes. Figure 2 illustrates the distribution of hostility levels in interstate disputes for the Early (1823–1913) and Modern (1914–2003) periods. The left panel shows that while disputes in the Modern period are less likely to escalate into full-scale wars, they are more likely to involve the use of force compared to those in the Early period. The right panel disaggregates this trend by distinguishing between disputes involving highly unequal combatants and those where military capacities were more balanced. Here, define a dispute as highly unequal if one side has at least five times the CINC score of its opponent[1]. The figure suggests that although highly unequal disputes are somewhat less likely to escalate into war, this difference is minor and becomes negligible in the Modern period.

Main Causes of Interstate Disputes Next, we examine whether the primary causes of interstate disputes have changed over time and whether these patterns differ for highly unequal disputes. Figure 3 presents the distribution of dispute motivations. The left panel shows that in both periods, foreign policy disagreements were the most common cause of disputes, followed by territorial disputes. This trend became even more pronounced in the Modern period.

The right panel disaggregates these trends by distinguishing between highly unequal disputes and others. In the Early period, territorial disputes were slightly more common

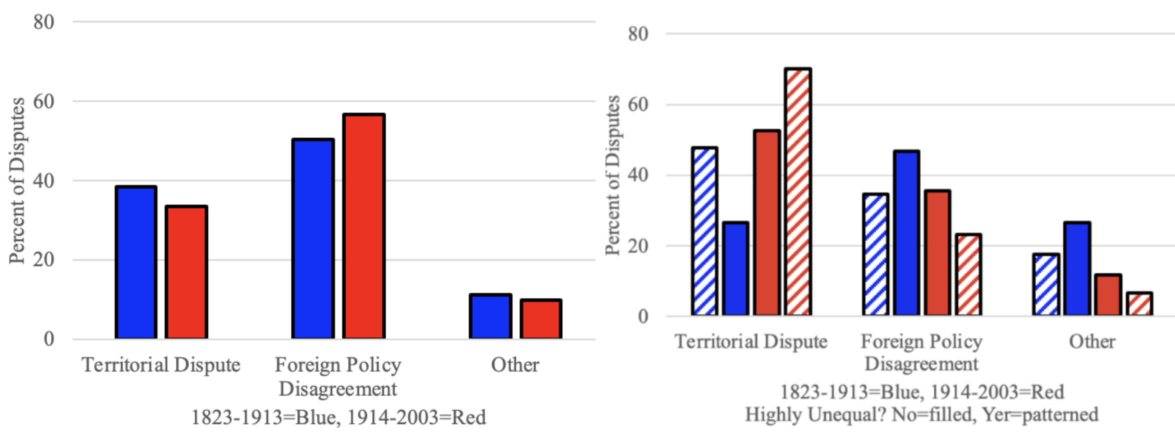
[1] The Composite Index of National Capacity (CINC), introduced by Singer (1987) and Singer et al. (1972), serves as a proxy for a nation's military capacity. It is calculated by averaging a country's global share of six war-relevant resources. While designed to measure a nation's ability to sustain armed conflict, the index has notable limitations. The most significant critique is that technological advancements alter the relationship between physical resources and military capacity over time. To address this issue, the model introduces an alternative instrument that estimates military capacity by isolating the common evolution of military technology and cyclical fluctuations in armament accumulation—factors likely influenced by expectations of war or continued conflict.



(a) Overall (b) Disaggregated

Figure 2: Distribution of interstate hostility levels in the Early and Modern periods, disaggregated by whether the disputes were highly unequal.

among highly unequal disputes, whereas disputes over foreign policy were more prevalent in more balanced conflicts. However, this pattern reverses in the Modern period, where highly unequal disputes are now more likely to be driven by foreign policy disagreements than territorial issues. This shift suggests that while highly unequal disputes have always had different underlying motivations, broader historical trends have influenced their evolution.

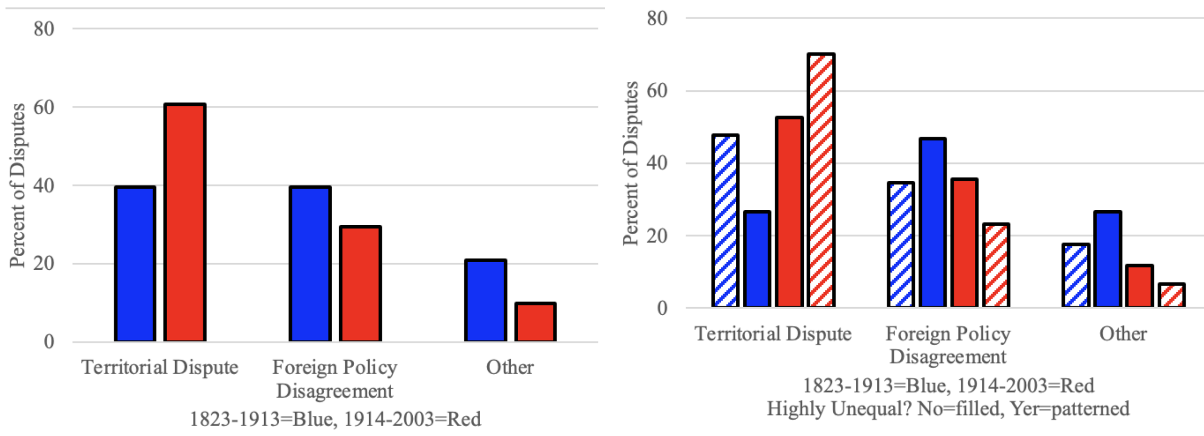


(a) Overall (b) Disaggregated

Figure 3: Distribution of main reasons for interstate disputes by period and by whether the dispute is highly unequal.

Do Certain Dispute Types Lead to War? A key limitation of analyzing dispute motivations is that the initial cause of a dispute may differ from the factors that ultimately drive it to escalate into war. Figure 4 explores this issue by examining the subset of disputes that escalated into full-scale wars.

The left panel shows that although foreign policy disagreements are the most common source of disputes, territorial disputes are disproportionately likely to escalate into war. The right panel further disaggregates this trend by comparing highly unequal disputes to others. The results suggest that the link between territorial disputes and war is even stronger among highly unequal conflicts. In contrast, in the Early period, wars between more evenly matched opponents were more likely to originate from foreign policy disagreements.



(a) Overall

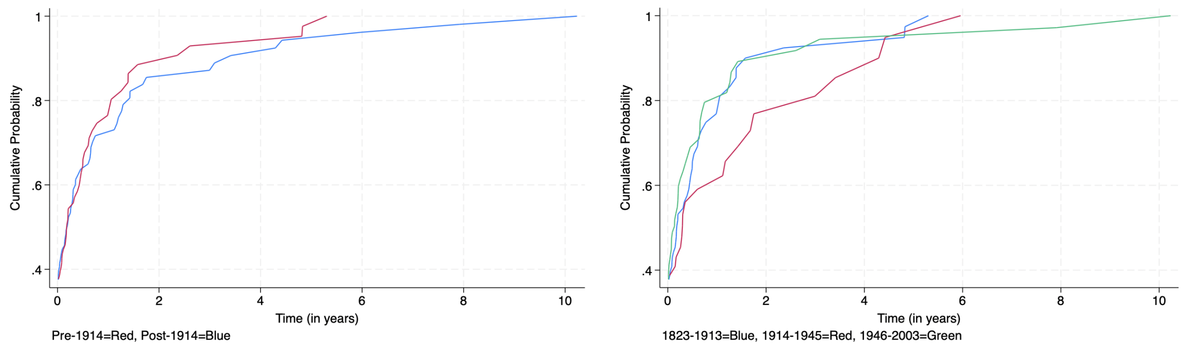
(b) Disaggregated

Figure 4: Distribution of main reasons for dispute by period and by whether the dispute is highly unequal, conditional on the dispute escalating into war.

5.4 Temporal Trends in War Duration

This section examines trends in war duration, focusing on how different characteristics of interstate conflicts are associated with their length. Rather than relying solely on summary statistics, the section presents cumulative distribution functions of war durations, offering a clearer comparison between periods and war characteristics.

Figure 5 plots the cumulative hazard rates for wars before and after 1914. The distributions suggest that while wars in both periods followed broadly similar patterns, the key difference is the greater frequency of wars lasting over three years in the post-1914 period. Further dividing the post-1914 period into 1914–1945 and 1946–2003 reveals a distinct pattern: the 1914–1945 period stands out, with wars tending to last longer than in both the earlier and later periods. This aligns with expectations given the geopolitical instability of the time. Once this period is accounted for, there are no major differences in war duration distributions between pre-1914 and post-1945 wars.



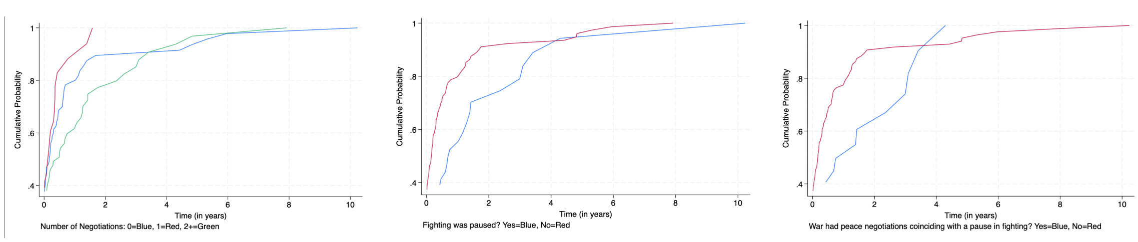
(a) Pre- and Post-1914

(b) Pre-1914, 1914–1945, 1946–2003

Figure 5: Cumulative hazard rates of war durations across periods.

Peace Negotiations and Pauses in Fighting A key question is whether peace negotiations, pauses in fighting, or their interaction influence war duration. The model predicts that pauses in fighting, particularly when accompanied by negotiations, should prolong conflict. Figure 6 presents descriptive evidence supporting this claim.

The first panel shows that wars with only a single negotiation tend to be shorter than those with multiple negotiations. This does not imply that negotiations cause wars to last longer but highlights an important correlation that warrants further analysis. The second panel shows that wars without pauses between major battles tend to be shorter than those with intermittent fighting, which is unsurprising given that the time between battles contributes to overall war duration. The third panel provides the most direct evidence supporting the model: wars where negotiations coincided with pauses in fighting lasted significantly longer, aside from a few outlier conflicts.



(a) 0, 1, or 2+ negotiations

(b) Fighting paused?

(c) Paused while negotiating?

Figure 6: Cumulative hazard rates by number of negotiations and pauses in fighting.

Survival Analysis with Cox Proportional Hazards Regression To formally test these relationships, Table 2 presents results from a Cox proportional hazards regression. The model examines how pauses in fighting, peace negotiations, and their interaction correlate with war duration.

The results indicate that wars with only peace negotiations (without pauses in fighting) do not exhibit statistically significant differences in duration. In contrast, wars with pauses in fighting (without negotiations) are associated with significantly lasting longer. The association is stronger when pauses and negotiations coincided, suggesting that such conflicts are particularly prolonged. However, the magnitudes of these effects do pale in comparison to the association between observing a decisive military victory and a war's termination.

Table 2: Cox Proportional Hazards Regression Results

Variable	**Hazard Ratio**
War had pauses in fighting or negotiations?	
No negotiations, did pause	0.332**
No pauses, did negotiate	0.836
Both pauses and negotiations	0.245***
Decisive Victory	
Decisive military victory	16.867**

Notes: Stars indicate statistical significance at $^{***}p < 0.01$, $^{**}p < 0.05$.

Taken together, these results support the model's prediction that pauses in fighting, especially when coinciding with negotiations, are associated with prolonged wars. However, this analysis does not establish causality. The subsequent section employs a panel instrumental variables (IV) approach to address potential endogeneity and provides a stronger causal argument.

5.5 Instrumental Variable Analysis: Integrating Technological Evolution in Military Capability

To address potential endogeneity concerns, this section introduces an instrumental variables (IV) approach within a panel regression framework. The section attempts to estimate the causal relation between a the time spent in fighting with formal peace negotiations coinciding with pauses in fighting. The analysis employs two distinct instruments, which are available starting in 1960 and 1970, respectively.

The first instrument, derived from Thomson et al. (2023), captures long-run trends in non-fatal social unrest as a proxy for the domestic political costs of war. Governments engaged in military conflicts often face internal dissent, which can limit their ability to sustain prolonged fighting. However, a major concern is that ongoing wars may trigger unrest, creating a potential endogeneity problem. To address this, we extract the long-term component of social unrest.

The second instrument, derived from Gannon (2023), estimates a nation's baseline military readiness by analyzing historical trends in military procurement and technological capacity. This measure captures a country's structural ability to attain decisive victories, independent of short-term shocks to military stockpiles or strategic wartime investments. To ensure that the instrument is exogenous to wartime mobilization, we isolate long-term trends in armament accumulation, removing short-term variations linked to pre-war military buildups.

5.5.1 Instrument 1: Long-Run Social Instability Trends

The first instrument leverages long-run domestic instability trends as a proxy for the political costs governments face when sustaining a war. Interstate wars divert government resources and attention away from internal affairs, often allowing social unrest to escalate. While civil wars and insurgencies can directly threaten regime survival, lower-intensity instability—such as protests, strikes, and political demonstrations—rarely forces a government to unilaterally withdraw from a conflict. However, persistent domestic unrest can preassure government officials to initiate negotiations and implement temporary pauses in fighting in order to shift attention towards domestic matters, particularly in democratic regimes. A key concern is whether ongoing wars themselves contribute to rising instability, which would violate the exclusion restriction. To ensure exogeneity, this instrument isolates the long-run component of domestic unrest, separating structural trends from short-term fluctuations driven by wartime events. We construct a time series $\{x_{it}\}$ measuring monthly non-fatal political disturbances in country i, sourced from Thomson et al. (2023), which records protests, strikes, and political demonstrations in 186 national capitals from 1960 to 2014. These events are then decomposed into a persistent long-run trend and short-term shocks:

$$x_{it} = \ell_{it} + \epsilon_{it} \tag{9}$$

where ℓ_{it} captures structural instability, while ϵ_{it} reflects temporary unrest triggered by war events, economic crises, or election cycles. This decomposition ensures that the instrument captures deep-rooted instability rather than immediate reactions to conflict.

The validity of this instrument hinges on the assumption that long-run instability affects war duration only through its influence on conflict intensity and ceasefire incentives, rather than through direct channels. One concern is reverse causality—if wars themselves increased instability trends, the instrument would be endogenous. However, empirical evidence suggests otherwise. Figure 7 shows that U.S. domestic instability was rising before the Vietnam War, suggesting that pre-existing political tensions, rather than war dynamics, drove unrest.

Instability peaked in 1968, before military de-escalation began, reinforcing the idea that the long-run trend was not merely a reflection of wartime disturbances. To further validate this, placebo tests examine whether instability trends predict the duration of militarized disputes that did not escalate into full-scale wars. If the instrument truly operates through bargaining incentives rather than directly affecting war length, it should not systematically predict the duration of non-war disputes.

Another concern is omitted variable bias. If domestic instability correlates with economic or institutional factors that independently influence war duration, the exclusion restriction would be violated. To address this, country fixed effects are included to control for time-invariant national characteristics, and decade fixed effects absorb global trends in conflict resolution. Additional robustness checks control for GDP per capita, democracy indices, and military spending to ensure that instability is not simply capturing broader political fragility.

To illustrate the exogeneity of long-run instability, Figure 7 plots trends ℓ_{it} for the United States (red) and North Vietnam (blue) during the Vietnam War. Several key patterns emerge. U.S. instability was increasing before military escalation, suggesting that pre-existing political tensions—rather than war dynamics—drove unrest. Instability peaked before military de-escalation in 1968, implying that it was not simply reacting to withdrawals. In contrast, domestic instability in North Vietnam remained consistently lower, reinforcing the idea that democratic regimes face greater internal pressures to negotiate.

Crucially, these patterns show that instability operates on a different temporal scale than conflict outcomes. It tends to shape early institutional responses—such as initiating peace talks or mandating mediation—but fades before the conflict's resolution. This timing mismatch reduces the likelihood that instability affects conflict duration directly, satisfying the exclusion restriction.

To generalize beyond Vietnam, Figure 8 examines two additional conflicts: the Iran-Iraq War and the Afghan-Soviet War. If instability were responding endogenously to war cycles, there would be sharp movements in trends around the start and end of conflicts. Instead, Iran experienced declining instability before the war, with unrest rising only during the Iranian Revolution—an event largely orthogonal to the war itself. Similarly, Afghanistan exhibited stable social unrest during peak fighting, with instability declining only after military withdrawal. These cases support the broader claim: long-run unrest influences intervention decisions but does not track conflict duration, mitigating concerns of reverse causality.

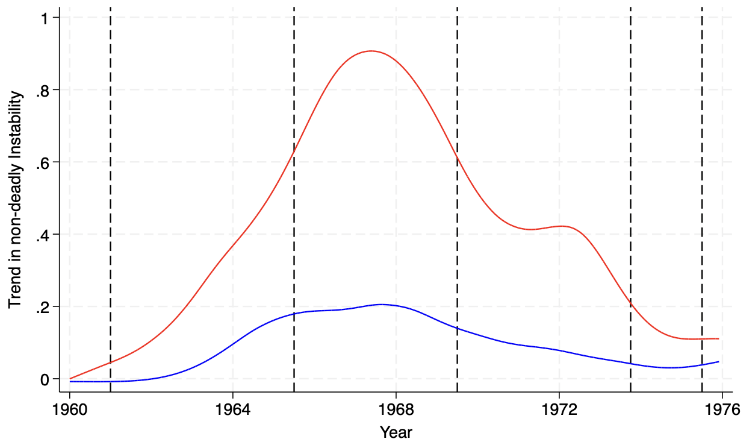
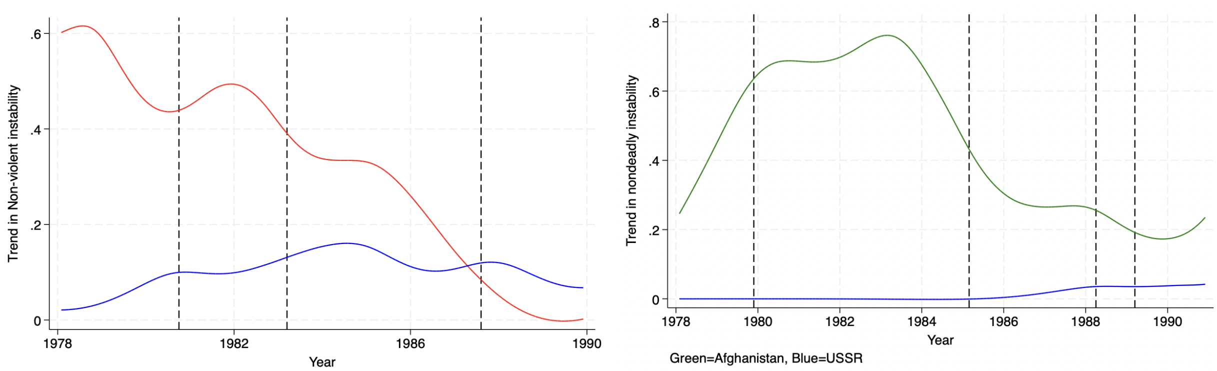


Figure 7: Trends ℓ_{it} for the U.S. (red) and North Vietnam (blue), alongside key dates of the Vietnam War.



(a) Iran-Iraq War (b) Afghan-Soviet War

Figure 8: Trends and key events for the Iran-Iraq and Afghan-Soviet Wars. Social instability trends appear uncorrelated with war cycles.

5.5.2 Instrument 2: Military Preparedness

The second instrument constructs a measure of military preparedness that isolates long-term structural capacity trends from short-term military fluctuations. Military capacity is a crucial determinant of bargaining leverage in wars, yet most existing metrics conflate long-run military capability with short-term mobilization decisions. To address this, we decompose military stock into three key components: a global trend reflecting technological development, a country-specific long-term trend, and short-term cyclical shocks.

We use data from Gannon (2023) on disaggregated counts of armaments across national armies from 1970 to 2014. Define the total military stock for country i at time t as y_{it} and decompose it as follows:

$$y_{it} = \tau_t \tau_{it} + c_{it} + c_t, \tag{10}$$

where τ_t represents global technological evolution, τ_{it} captures country-specific long-term military capacity, c_{it} denotes cyclical shocks specific to country i, and c_t represents global cyclical factors.

To construct a measure of relative military preparedness, we define r_{it} as:

$$r_{it} \equiv \frac{\hat{\tau}_{it}}{\bar{\tau}_t} \approx \frac{\tau_{it}}{E_t[\tau_{it}]},$$

where $\bar{\tau}_t$ is the cross-country average trend in capacity. This measure reflects a country's relative standing in military preparedness over time, distinguishing between structural capability and short-run mobilization decisions.

Exclusion Restriction and Endogeneity Concerns For r_{it} to serve as a valid instrument, it must influence war duration only through its impact on bargaining power and not through omitted variables such as geopolitical shifts, economic downturns, or endogenous war strategies. One potential concern is that military preparedness may be shaped by strategic geopolitical rivalries rather than exogenous technological advancements. However, the absence of sharp trend shifts around major conflicts (Figure 10) suggests that r_{it} is not systematically driven by endogenous wartime mobilization.

Another concern is reverse causality—if governments adjust military capacity in response to war onset, r_{it} would be endogenous. To rule this out, we estimate models using r_{it} lagged by 5 and 10 years. The results confirm that military preparedness trends are predetermined relative to war onset, reinforcing the exogeneity of our instrument.

Validation and Empirical Evidence We compare r_{it} with the Composite Index of National Capacity (CINC) from Singer (1987), which averages a country's global share of resources such as population, military spending, and industrial output. While CINC captures broad military potential, it fails to account for technological progress and efficiency gains in war production. Figure 9 shows that the correlation between CINC and r_{it} was strong before 1990 but weakened thereafter, consistent with the notion that CINC overstates the military capacity of countries that failed to modernize relative to their actual war readiness.

To further validate r_{it}, Figure 10 presents trends across key geopolitical contexts. In Panel (a), the sharp decline in Russia's capacity after the Soviet collapse aligns with historical accounts of military contraction. In contrast, the U.S. military buildup post-Vietnam is visible in the upward trend in r_{it}, followed by stabilization after the 1990s. Meanwhile, China's stable trend despite involvement in conflicts such as the Sino-Vietnamese War suggests that r_{it} captures long-run industrial capacity rather than short-term wartime fluctuations.

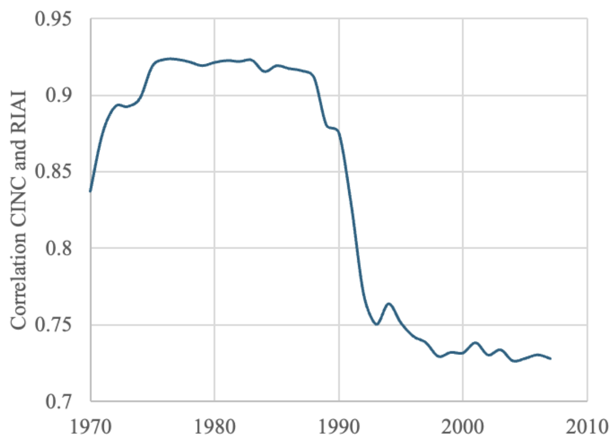
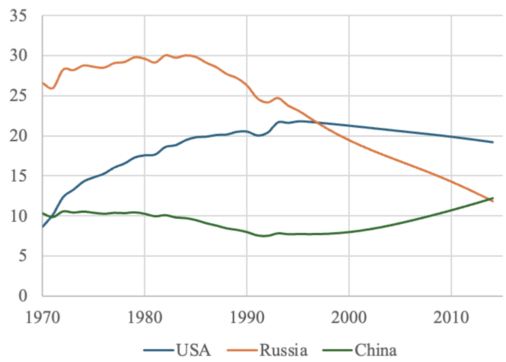


Figure 9: Yearly correlation between CINC and r_{it} from 1970 to 2007.

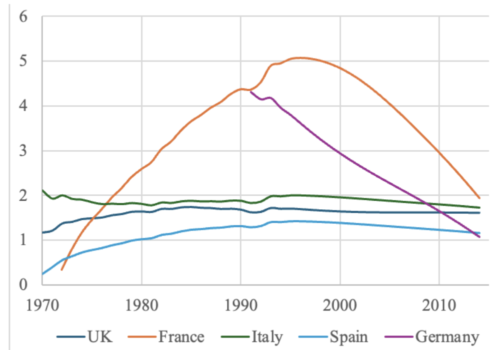
Importantly, many low-preparedness countries—such as Iran and North Korea—engage in short but intense conflicts that resolve quickly, particularly in the absence of third-party mediation. This cross-sectional variation weakens concerns that r_{it} directly prolongs disputes. Instead, the evidence suggests that preparedness influences the *likelihood* of institutional intervention (e.g., ceasefires or external pressure), rather than independently shaping conflict duration. This supports the exclusion restriction required for instrumental validity.

Instrument Validity and Empirical Validation Both instruments are designed to exploit variation that is plausibly exogenous to short-run conflict dynamics. To test the exclusion restriction directly, we conduct placebo regressions using a sample of militarized interstate disputes (MIDs) that did not escalate to full-scale war. In these cases, institutional negotiation mechanisms such as ceasefires or formal mediation were rarely triggered, allowing us to isolate whether the instruments have independent predictive power for dispute duration outside the context of wartime bargaining. As shown in Table 4, neither instrument significantly predicts the length of these non-war disputes. This suggests that their effect on duration operates primarily through the institutional mechanisms present in full-scale wars.

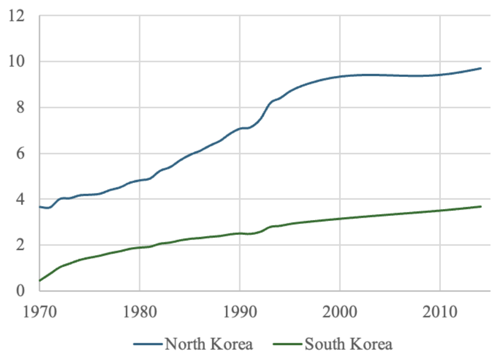
In addition, Table5 presents pairwise correlations between the instruments, dispute duration, and the levels of hostility. The correlations are uniformly small and statistically insignificant, further supporting the claim that the instruments are not proxies for conflict severity or unobserved aggression. Taken together, these results bolster the argument that the instruments affect war duration primarily through their influence on institutional incentives for conflict management—consistent with the identifying assumptions of the IV strategy.



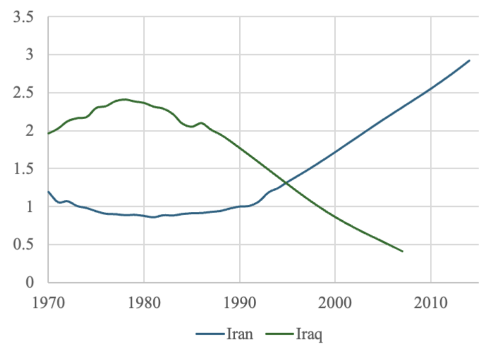
(a) US vs. Russia vs. China



(b) Western Europe



(c) Korea



(d) Iran vs. Iraq

Figure 10: Relative trends in r_{it} across key conflicts, showing that the measure follows long-term military evolution rather than short-term war mobilization.

5.5.3 Regression Results

Table 3 presents the results from the panel regressions, which examine the relationship between peace negotiations, pauses in fighting, and the duration of conflicts. The dependent variables measure (i) the share of time spent actively fighting that has already elapsed and (ii) the share of total war duration that has already passed. The key distinction between these measures is that the first accounts only for days when major battles occurred, relative to the total number of battle days, while the second considers all days in which the war was ongoing, irrespective of active combat. These variables allow for the assessment of how formal negotiations and temporary pauses in combat influence war duration, both individually and in interaction with one another.

In the baseline (non-IV) regressions, peace negotiations are generally associated with shorter wars, as indicated by the positive and statistically significant coefficients across most historical periods. Similarly, pauses in fighting also correlate with reductions in war duration. However, the interaction term between negotiations and pauses in fighting is consistently negative and significant, suggesting that when peace talks occur during a temporary halt in combat, conflicts tend to last longer overall.

To address potential endogeneity, we estimate a panel IV regression using the instrumental variables $\{\ell_{it}, r_{it}\}$, which capture long-run social instability trends and structural military preparedness. Since data on r_{it} is unavailable before 1969, we set $r_{it} = 0$ for earlier years to maintain a balanced panel.

The IV estimates strongly support the model's predictions. First, the instruments pass standard weak IV robustness tests, confirming their strength and validity. The results indicate that peace negotiations and temporary pauses in fighting independently reduce the share of time spent in combat, even if no formal agreement is reached. However, consistent with the theoretical model, the interaction term between negotiations and pauses remains strongly negative and statistically significant, suggesting that when peace talks coincide with temporary lulls in fighting, wars are prolonged rather than shortened. The magnitude of this effect is large enough to offset the individual benefits of either peace talks or pauses in fighting, implying that such negotiations may inadvertently reduce incentives for resolution by allowing combatants to regroup and reassess their strategic positions.

While dropping the long-run instability instrument reduces precision and renders the effect statistically insignificant, the direction and magnitude of the estimates remain stable. This suggests that the full-sample result is not driven by a single instrument but reflects the joint predictive content of two plausibly exogenous measures of intervention propensity. All specifications pass weak-IV tests.

6 Discussion

6.1 Broader Implications and Applications

Although the empirical focus is on interstate war, the theoretical insights extend naturally to a broad class of strategic disputes—including labor negotiations, corporate takeovers, legislative standoffs, and trade conflicts. A central implication is that bargaining power is not static. Rather, it evolves endogenously through each party's ability to impose and absorb costs over time. This perspective challenges conventional views that power asymmetries mechanically determine outcomes and instead emphasizes that long-run endurance can outweigh short-run dominance.

The model also reframes how we understand the strategic consequences of shifting bargaining conditions. In standard models, power shifts—whether due to battlefield gains, political developments, or institutional changes—tend to generate rapid concessions and early resolution. In contrast, the dynamics here suggest a more subtle mechanism. When

Table 3: IV regression results using the new instrument.

	Share of Fighting Already Done						Share of Time Spent in War Already Done					
VARIABLES	1823-2003	1823-1913	1914-1945	1946-2003	1960-2003	IV: 1960-2003	1823-2003	1823-1913	1914-1945	1946-2003	1960-2003	IV: 1960-2003
Negotiation	0.058**	0.175**	0.176**	-0.026*	0.035**	4.260**	0.054**	0.176**	0.178**	-0.032*	0.025**	3.829**
Interaction	-0.056**	-0.068**	-0.064**	-0.050**	-0.055**	-5.085**	-0.056**	-0.047**	-0.056**	-0.053**	-0.056**	-5.132**
Pause	0.008**	0.006**	0.009**	-0.007**	0.009**	1.216**	0.005**	0.009**	0.009**	-0.006**	0.011**	0.964**
Day number	0.001***	0.001***	0.001***	0.000	0.001**	0.000	0.001***	0.001***	0.001***	0.000	0.001**	0.000
A battle has been won	-0.005	-0.028	0.000	-0.083**	-0.166**	-0.543*	-0.012	-0.048**	-0.022	-0.086**	-0.161**	-0.592*
UN Sec. Council Member involved	0.011***	-	-	-0.013*	-0.169**	-0.255**	0.012***	-	-	-0.015**	-0.106**	-0.221**
Nuclear Capability	0.200**	-	-	-0.381**	-0.588**	-	0.216**	-	has bee-	-0.416**	-0.606**	-
Number of Allies	0.002**	0.004***	0.004***	0.004**	0.579	-0.001**	-0.001**	-0.001**	-0.001*	0.007	-	
Relative energy production	0.209**	-17.612**	-13.055**	10.067**	0.761	-9.440**	0.256**	-17.519**	8.511**	0.396	-7.401**	
Relative Military Personnel	0.043**	1.471**	0.000	0.380**	0.825**	1.471**	0.049**	1.557**	0.430**	0.554**	-	
Relative Military Expenditure	-0.311**	-0.678**	-0.599**	-0.191**	-0.220	-0.858**	-0.291**	-0.615**	-0.545**	-0.291**	-0.430**	-
Relative Urban Population	0.061**	2.331**	2.819**	17.696**	13.396**	45.911**	0.222**	2.368**	2.795**	13.071**	42.482**	-
Relative Total Population	-1.738**	-6.315**	-13.868**	-54.571**	40.283**	-57.077**	-1.580**	-7.198**	-10.301**	-49.968**	-59.995**	-
Relative Iron and Steel Production	1.083**	18.607**	16.871**	-3.854**	3.031**	7.161**	0.956**	19.681**	18.164**	-3.314**	4.847**	-
Year	-0.033**	-0.013**	-0.025**	-0.121**	-0.089**	-0.008	-0.014**	0.021**	0.010**	-0.112**	-0.082**	0.001
Constant	62.968***	23.907***	45.227***	239.476***	175.639***	10.729	27.415***	-38.364***	-19.761***	223.631***	161.064***	-5.256
Observations	36,848	12,524	12,388	13,122	10,845	14,887	36,848	12,524	12,388	13,122	10,845	14,887
R-squared	0.653	0.717	0.722	0.708	0.728	-	0.665	0.740	0.743	0.735	0.759	-
Number of Wars	92	38	37	35	27	22	92	38	37	35	27	22
Weak IV Robustness Tests												
CLR Statistic	-	-	-	-	-	7058.14	-	-	-	-	-	6379.76
K Test (Chi2)	-	-	-	-	-	4136.39	-	-	-	-	-	3926.25

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

power shifts are anticipated, parties expecting to lose leverage often slow their rate of concession, hoping that external conditions will shift in their favor. If no immediate settlement is reached, negotiations re-enter a war-of-attrition phase, characterized by reputational persistence and strategic delay. Here, delay is not simply a byproduct of information asymmetries, but a forward-looking strategy in an evolving environment.

These dynamics are observable across diverse domains. Unions may delay resolution in anticipation of favorable regulatory shifts. Firms targeted for acquisition might stall while awaiting antitrust decisions. Trade partners may drag out disputes expecting macroeconomic trends to tilt bargaining leverage. In each case, the model emphasizes that concessions depend not only on current power, but on how actors anticipate future shifts in bargaining conditions.

6.2 The Strategic Effects of Pacifying Policies

One of the model's more counterintuitive predictions is that pacifying policies—such as ceasefires, mandated mediation, or strike bans—can prolong disputes. Though these interventions are intended to reduce conflict intensity and accelerate settlement, they often have the opposite effect: by lowering the immediate cost of delay, they give parties greater incentive to wait rather than concede.

The logic unfolds in three steps. First, pacifying policies reduce the likelihood of immediate settlement by dulling the urgency to reach agreement. Second, they slow the rate of incremental concessions, as the short-run cost of holding out diminishes. Third, they weaken the effect of power shifts: if a party expects a future advantage to materialize, a pacified environment allows it to wait longer before making concessions. Together, these mechanisms help explain why ceasefires often extend wars, why mediation can stall labor negotiations, and why procedural rules sometimes entrench political gridlock.

Importantly, these inefficiencies persist even in the absence of classical frictions. In the model, when reputational uncertainty and commitment problems vanish, disputes still endure—but only if pacifying policies are in place. Without such interventions, the model predicts immediate resolution in the frictionless limit. This highlights that the institutional environment—not just the strategic structure of the game—can sustain inefficiency.

These results raise sharp policy trade-offs. Ex ante, pacifying policies may encourage actors to initiate disputes they would otherwise avoid. Ex post, they are difficult to unwind due to humanitarian appeal or institutional inertia. This tension mirrors time inconsistency problems in financial regulation: interventions designed to stabilize crises may inadvertently incentivize risk-taking, thereby increasing the frequency and duration of crises over time.

6.3 Empirical Insights and Methodological Considerations

The empirical analysis supports the model's predictions. Using a 200-year dataset of interstate wars, the evidence shows that ceasefires often prolong conflict, weaker parties extract concessions despite material disadvantage, and outcomes are shaped more by strategic endurance than by raw military capacity. These patterns reinforce the model's central claim: that the path of conflict is driven by evolving incentives, not just static power.

To estimate the causal effects of pacifying policies, the paper employs a panel instrumental variables strategy leveraging two slow-moving sources of exogenous variation: domestic instability and long-run military preparedness. The first instrument isolates structural patterns of unrest and institutional fragility. These factors, which evolve slowly over time, affect governments' willingness to sustain conflict but are unlikely to respond to short-term military events. The second instrument measures deep military capacity—rooted in procurement patterns, technological infrastructure, and industrial capability—which shapes bargaining leverage independently of wartime mobilization.

While these instruments mitigate endogeneity concerns, they do not capture all aspects of bargaining dynamics. Ephemeral factors—such as political signaling, tactical surprises, or backchannel diplomacy—are harder to measure and may play important roles in short-term decisions. Moreover, structural instruments better explain protracted disputes than rapid escalations or sudden settlements.

Future work can extend these insights using high-frequency datasets, case studies, or quasi-experimental variation in policy interventions. Real-time data on conflict intensity, negotiation timelines, or institutional interventions could clarify short-run strategic behavior. Experimental approaches might further isolate the mechanisms through which well-intentioned pacifying policies prolong conflict.

7 Conclusion

Persistent disputes—in war, labor negotiations, or corporate and legal settings—often defy standard bargaining predictions. Weaker parties extract concessions despite asymmetric power, and institutional interventions such as ceasefires or arbitration frequently fail to accelerate resolution. This paper offers a formal explanation for these outcomes through a dynamic reputational bargaining model with power shifts and indivisible stakes.

The model departs from canonical results by showing that inefficiencies can persist even as reputational frictions vanish—*but only when pacifying policies are imposed.* When interventions reduce the short-run cost of conflict, they weaken incentives to settle, enabling strategic delay. Delay arises not from misperception or incomplete information, but from a credible expectation of future advantage under softened conditions. The result is a fundamental time inconsistency in institutional design: the very tools meant to resolve disputes can entrench them.

Empirical analysis of two centuries of interstate wars supports the model's predictions. Ceasefires and other de-escalatory policies are associated with longer active conflicts. Conflict outcomes are shaped less by static power asymmetries than by the capacity to endure over time. The empirical results extend to a broader set of interstate disputes, underscoring the generality of the mechanism.

These findings have clear policy implications. In military settings, ceasefires may serve as opportunities to regroup, not pathways to peace. In legal and corporate contexts, pauses in litigation or arbitration can delay, rather than hasten, settlement. In labor negotiations, banning strikes may slow resolution by removing pressure to compromise. Across these domains, well-meaning interventions can backfire—not because they fail to reduce violence or cost, but because they shift incentives in ways that make conflict persist. These patterns highlight a broader design challenge: interventions that appear optimal in the heat of conflict may be suboptimal in expectation, reinforcing the time-inconsistency at the heart of many conflict-resolution institutions.

More broadly, the paper contributes a unified framework for understanding delay in institutionalized bargaining. It highlights the unintended consequences of dispute-mitigation tools and identifies conditions under which inefficiencies arise even in frictionless environments. Future work could explore how institutional design might mitigate this tradeoff—preserving peacekeeping or mediation efforts while restoring incentives for timely resolution. This tradeoff underscores a broader dilemma in institutional design. Interventions that reduce the short-run risks of conflict—by muting confrontation, suspending strikes, or mandating negotiation—can unintentionally create long-run incentives for strategic delay. Policymakers face a difficult choice: act early and risk entrenching inefficiency, or withhold support and risk escalation. Optimal institutional design may thus require strategic restraint: crafting mechanisms that selectively tolerate friction or confrontation to preserve pressure for resolution. Determining when to intervene—and when not to—remains a central challenge across domains from diplomacy and labor law to litigation and corporate governance.

References

Abreu, D. and F. Gul (2000). Bargaining and reputation. *Econometrica 68*(1), 85—117.

Alesina, A. and G. Tabellini (2007). Bureaucrats or politicians? part i: A single policy task. *American Economic Review 97*(1), 169–179.

Beardsley, K. (2011). *The Mediation Dilemma*. Ithaca, NY: Cornell University Press.

Davion, I. (2020). Encyclopédie d'histoire numérique de l'europe: Ending wars. Retrieved from https://ehne.fr/en/encyclopedia/themes/wars-and-memories/victors-and-vanquished-in-europe/ending-war.

Ekmekci, M. and H. Zhang (2024). Reputational bargaining with ultimatum opportunities. *Review of Economic Studies 0*(0), Forthcoming.

Fanning, J. (2021). Mediation in reputational bargaining. *American Economic Review 111*(8), 2444–2472.

Fearon, J. (1995). Strategicist explanations for war. *International Organization 43*(3), 379—414.

Fearon, J. (2013). Fighting rather than bargaining. *Annual Meetings of the American Political Science Association 4*(0), 1—46.

Filson, D. and S. Werner (2002). A bargaining model of war and peace: Anticipating the onset, duration, and outcome of war. *American Journal of Political Science 46*(4), 819—837.

Gannon, A. (2023). Planes, trains, and armored mobiles: Introducing a dataset of the global distribution of military capabilities (rdmc). *International Studies Quarterly 67*(4), 1–12.

Jackson, M. and M. Morelli (2009). The reasons for wars – an updated survey. In *Handbook on the Political Economy of War*.

Krainin, C., C. Thomas, and T. Wiseman (2020). Rational quagmires: Attrition, learning, and war. *Quarterly Journal of Political Science 15*(3), 369—400.

Leventoglu, B. and B. Stanchev (2007). The armed peace: A punctuated equilibrium theory of war. *American Journal of Political Science 51*(4), 755—771.

Lyall, J. (2020). *Divided Armies: Inequality and Battlefield Performance in Modern War.* Princeton, New Jersey: Princeton University Press.

Min, E. (2020). Talking while fighting: Understanding the role of wartime negotiation. *International Organization 74*(3), 610—632.

Min, E. (2021). Interstate war battle dataset (version 1.0). Retrieved dataset from : https://doi.org/10.7910/DVN/KLQFAP, Harvard Dataverse, V1, UNF:6:Rslo7+WxkM0f40mVY23xfQ== [fileUNF].

Moffitt, M. L. (2013). Mediating corporate governance conflicts and disputes. Corporate Governance Focus Series 4, International Finance Corporation (IFC).

Powell, R. (2004). Bargaining and learning while fighting. *American Journal of Political Science 48*(2), 344—361.

Sarkees, M. and F. Wynman (2010). Resort to war: 1816 – 2007. Washington DC: CQ Press. Extracted from https://correlatesofwar.org/data-sets/cow-war/.

Schelling, T. (1967). Arms and influence. *American Political Science Review 61*(3), 766—767.

Singer, D. (1987). Reconstructing the correlates of war dataset on material capabilities of states, 1816-1985. *International Interactions 14*(0), 115–132.

Singer, D., S. Bremet, and J. Stockey (1972). Apability distribution, uncertainty, and major power war, 1820-1965. *in Bruce Russett (ed) Peace, War, and Numbers Sage ?*(?), 19–48.

Thomson, H., K. Bahgat, H. Urdal, and H. Buhaug (2023). Urban social disorder 3.0: A global, city-level event dataset of political mobilization and disorder. *Journal of Peace Research 60*(3), 521–531.

Tirole, J. (2015). Societal commitment devices. *Econometrica 83*(3), 1457–1501.

UN, U. N. (2022). Guidance on mediation of ceasefires. Retrived https://peacemaker.un.org/sites/peacemaker.un.org/files/Ceasefire-Guidance-2022.pdf.

A Derivations

A.1 Derivation of Equilibrium Payoffs and Concession Rates

Let $W_{i\ell t}$ denote player i's expected discounted payoff at a strategic, stable time $t > 0$, conditional on being in state ℓ. Given that the state process follows a continuous-time Markov chain and that behavior evolves independently across types, $W_{i\ell t}$ satisfies the following Feynman-Kac equation:

$$rW_{i\ell t} = -(\bar{c} + \psi_t c_{j\ell}) + \mu_{jt}\phi(-B_{i\ell} - W_{i\ell t}) + C_{j\ell t}(1 - W_{i\ell t})$$
$$+\psi_t\nu_{i\ell}(1-W_{i\ell t})+\psi_t\nu_{j\ell}(-W_{i\ell t})+\sum_{\ell'\neq\ell}\lambda_{t\ell\ell'}\left[(1 - q_{j\ell\ell' t})(1 - W_{i\ell t}) + q_{j\ell\ell' t}(W_{i\ell' t^+} - W_{i\ell t})\right]+\dot{W}_{i\ell t}. \tag{11}$$

At strategic, stable times, Lemma 2 implies that players are indifferent between conceding and continuing the dispute. Thus, we can set $W_{i\ell t} = \dot{W}_{i\ell t} = 0$. Define the total concession rate:

$$K_{i\ell t} \equiv C_{i\ell t} + \sum_{\ell'\neq\ell}(1 - q_{i\ell\ell' t})\lambda_{t\ell\ell'}. \tag{12}$$

Substituting into Equation (11), we obtain:

$$K_{i\ell t} = \bar{c} + \mu_{jt}\phi B_{i\ell} + \psi_t(c_{i\ell} - \nu_{j\ell}). \tag{13}$$

This concludes the derivation of the equilibrium concession rate as presented in the main text.

A.2 Derivation of Concession Probabilities and Belief Convergence

Let t denote a moment of state transition. As established in Lemma 2, at most one player concedes discontinuously at t. Without loss of generality, suppose player i is the conceding party. We seek to determine the concession probability $1 - q_{it}$, and the belief convergence time T_t that follows.

Define T_{it} as the time required for player j's belief about i's obstinacy to converge to certainty. The belief process satisfies the following integral condition:

$$0 = \ln \mu_{it^-} + \int_0^{T_{it}} \left(C_{i\ell t+s} - \phi(1 - \mu_{it+s})\right) \mathrm{d}s, \tag{14}$$

where μ_{it^-} is the belief about player i's obstinacy just prior to time t. The belief convergence time for the system is then defined as $T_t = \min_i T_{it}$.

Suppose player i was expected to concede at time t but does not. Bayes' rule then updates player j's belief, yielding:

$$\ln \mu_{it} = \ln \mu_{it^-} - \ln q_{it}.$$

To ensure convergence to certainty by time $t + T_{it}$, the updated belief must also satisfy the integral condition:

$$0 = \ln \mu_{it^-} - \ln q_{it} + \int_0^{T_{it}} \left(C_{i\ell t+s} - \phi(1 - \mu_{it+s})\right) \mathrm{d}s. \tag{15}$$

Combining Equation (14) for player j and Equation (15) for player i, we obtain the key equilibrium identity:

$$q_{it}\mu_{jt^-} = \mu_{it^-} \exp\left\{ \int_0^{T_{it}} \left[C_{i\ell t+s} - C_{j\ell t+s} + \phi(\mu_{it+s} - \mu_{jt+s})\right] \mathrm{d}s \right\}. \tag{16}$$

This condition expresses the equilibrium concession probability in terms of observable reputational dynamics and expected concession behavior. It captures the idea that if player i's reputation converges more slowly than player j's, they must compensate by conceding with higher probability. Conversely, if player j is expected to concede more rapidly over the interval $[0, T_t]$, the burden may fall on them instead. Since beliefs evolve endogenously with player behavior, the identity of the conceding party can shift depending on when the state transition occurs and how reputations have developed up to that point. The recursive structure of this logic is what underpins the broader delay result and equilibrium construction.

A.3 Backward Induction

To complete the equilibrium characterization, we derive closed-form expressions for the full set of equilibrium objects: $\left\{\{C_{i\ell t},\, q_{i\ell\ell' t},\, \mu_{it}\}_{i=1}^2,\, T_t\right\}$. Although previous results provide functional relationships between these objects, full determination requires a recursive approach. We proceed by backward induction on the state space.

Step 1: Terminal State $\ell = n$. Suppose the system reaches terminal state $\ell = n$ at time t, with beliefs $\max_i \mu_{it^-} < 1$. Since no further state transitions are possible, simplifies to:

$$C_{int+s} = \bar{c} + \mu_{it+s}\phi B_{jn} + \psi_{t+s}(c_{in} - \nu_{jn}). \tag{17}$$

This closed-form allows us to directly compute belief dynamics using equations 12 and (3). Combining with Equations (15) and (16), we can fully characterize $q_{i\ell nt}$ and T_t for any prior state $\ell < n$, conditional on pre-jump beliefs $\{\mu_{it^-}\}_{i=1}^2$. Importantly, the jump destination $\ell' = n$ determines these dynamics; the originating state ℓ is irrelevant. Thus, we write $q_{i\ell nt} = q_{int}(\mu_{it^-}, \mu_{jt^-})$ for all $\ell < n$.

Step 2: Inductive Step. Assume equilibrium quantities—$C_{i\ell t}$, $q_{i\ell\ell' t}$, T_t, $\dot{\mu}_{it}/\mu_{it}$—have been derived for all $\ell' = m, \ldots, n$. Then, for state $\ell = m-1$, Equation 12 becomes:

$$C_{im-1\,t} = \bar{c} + \mu_{it}\phi B_{jm-1} + \psi_t(c_{im-1} - \nu_{jm-1}) - \sum_{\ell'=m}^{n}(1 - q_{i\ell' t})\lambda_{t\,m-1\,\ell'}. \tag{18}$$

Because all $q_{i\ell' t}$ for $\ell' \geq m$ are known from prior steps, this equation uniquely determines $C_{im-1\,t}$. Then, applying Equation (3) gives $\dot{\mu}_{it}/\mu_{it}$, and the pair $(\mu_{it}, C_{im-1\,t})$ allows us to solve for $q_{i\ell\,m-1\,t}$ and T_t for all $\ell < m-1$, again via Equations (15) and (16).

As before, we can suppress the origin state and write $q_{i\ell\,m-1\,t} = q_{i\,m-1\,t}(\mu_{it^-}, \mu_{jt^-})$. Iterating this argument from $\ell = n$ down to $\ell = 1$ yields a complete characterization of all equilibrium quantities. This insight is described below.

Corollary 7. *Suppose a state transition occurs at time t, from ℓ to $\ell' > \ell$, and current beliefs are $\{\mu_{it^-}\}_{i=1}^2$. Then, the probability that player i concedes discontinuously at time t depends only on μ_{it^-}, μ_{jt^-}, ℓ', and t—not on the prior state ℓ.*

A.4 Derivation of Lemma 6.

Consider the special case derived in the sub-setion presenting Example 2. Since the state does not transition, all concessions at strategic, stable times $t > 0$ occur gradually. The equilibrium conditions imply that:

$$C_{i1t} = K_{i1t} = \bar{c} + \mu_{i1t}\phi B_{j1}, \quad C_{j1t} = K_{j1t} = \bar{c} + \mu_{j1t}\phi B_{i1}.$$

Equation (3) further implies that the evolution of beliefs satisfies:

$$\frac{\dot{\mu}_{i1t}}{\mu_{i1t}} = (\bar{c} - \phi) + \mu_{i1t}\phi(1 + B_{j1}), \quad \frac{\dot{\mu}_{j1t}}{\mu_{j1t}} = (\bar{c} - \phi) + \mu_{j1t}\phi(1 + B_{i1}).$$

Defining $g_{i1t} \equiv 1/\mu_{i1t}$ and $g_{j1t} \equiv 1/\mu_{j1t}$, and rewriting the above expressions obtains:

$$\frac{\mathrm{d}}{\mathrm{dt}}\left[e^{(\bar{c}-\phi)t}g_{i1t}\right] = -\phi(1 + B_{j1})e^{(\bar{c}-\phi)t}, \quad \frac{\mathrm{d}}{\mathrm{dt}}\left[e^{(\bar{c}-\phi)t}g_{j1t}\right] = -\phi(1 + B_{i1})e^{(\bar{c}-\phi)t}.$$

Integrating both expressions and applying the initial conditions g_{i10^+} and g_{j10^+}obtains:

$$g_{i1t} = e^{-(\bar{c}-\phi)t}g_{i10^+} - \phi\left(\frac{1 + B_{j1}}{\bar{c} - \phi}\right)(1 - e^{-(\bar{c}-\phi)t}),$$

$$g_{j1t} = e^{-(\bar{c}-\phi)t}g_{j10^+} - \phi\left(\frac{1 + B_{i1}}{\bar{c} - \phi}\right)(1 - e^{-(\bar{c}-\phi)t}). \tag{19}$$

Since $B_{i1} > B_{j1}$, player j must concede at time 0 with a positive probability. Otherwise, player i would concede faster, and their reputation for being obstinate would deteriorate more quickly than their opponent's. Thus, setting $g_{i10^+} = 1/\mu$ and $g_{j10^+} = q_0/\mu$, where q_0 represents the probability that player j does not concede immediately.

Now, let T denote the earliest time at which beliefs converge to 1. From Equation 19, one can derive:

$$(\bar{c} - \phi)T(\mu) = \ln \frac{1/\mu + \phi\frac{1+B_{j1}}{\bar{c}-\phi}}{1 + \phi\frac{1+B_{j1}}{\bar{c}-\phi}}.$$

Substituting T into Equation 19 at time T obtains:

$$1 + \phi\left(\frac{1 + B_{i1}}{\bar{c} - \phi}\right) = \left[\frac{q_0}{\mu} + \left(\frac{1 + B_{i1}}{\bar{c} - \phi}\right)\right]\left[\frac{1/\mu + \phi\frac{1+B_{j1}}{\bar{c}-\phi}}{1 + \phi\frac{1+B_{j1}}{\bar{c}-\phi}}\right]^{-1}.$$

Solving for q_0, we find:

$$q(\mu, \phi) = \left[1 + \mu\phi\frac{1 + B_{j1}}{\bar{c} - \phi}\right]\frac{1 + \phi\frac{1+B_{i1}}{\bar{c}-\phi}}{1 + \phi\frac{1+B_{j1}}{\bar{c}-\phi}} - \mu\frac{1 + B_{i1}}{\bar{c} - \phi}.$$

Taking the limit as $(\mu, \phi) \to 0$, in any order, obtains $q(\mu, \phi) \to 1$, namely, the probability of an initial concession converges to 0. This concludes the desired derivation.

B Placebo Test

This appendix presents a placebo test that probes the exclusion restriction underlying the paper's instrumental variables strategy. The instruments—initial latent military capacity and long-run social instability—are theorized to affect conflict duration only through their influence on negotiation dynamics during war. If, instead, these variables systematically affected the duration of disputes more generally, this would cast doubt on the identification strategy. To investigate this concern, we examine disputes that did not escalate into war. In such cases, the notion of a negotiated ceasefire or peace settlement is not applicable. Therefore, if the instruments were found to significantly predict dispute duration in this subsample, it would suggest a violation of the exclusion restriction.

Table 4 presents the regression results, where dispute duration among non-war, interstate conflicts serves as the dependent variable, and the instruments' initial values are included as independent variables. The estimated coefficients are not statistically significant, and the effect sizes are generally small, suggesting little to no relationship between the instruments and dispute duration in this context.

Table 1: Placebo Tests: Effect of Instruments on Duration of Non-War Interstate Disputes (Post-1970)

	Dispute Duration	
(lr)2-3	**All States**	**Excl. Cold War Powers**
Instigator's initial r_{it}	-4.23	6.91
	(3.50)	(11.77)
Instigator's initial social instability	-6.42	-10.08
	(33.77)	(34.61)
Target's initial r_{it}	-0.77	3.03
	(2.75)	(10.99)
Target's initial social instability	11.69	-9.97
	(95.64)	(102.07)
Observations	751	598
R-squared	0.071	0.075
Adjusted R-squared	0.038	0.033
F-statistic	2.12	1.78

Notes: This table reports placebo regressions testing whether the main IV instruments predict the duration of non-war interstate disputes (i.e., MIDs that did not escalate to war). Standard errors in parentheses. None of the instrument coefficients are statistically significant. All models include controls for instigator and target characteristics, issue type, and hostility levels. The second column excludes Cold War powers (U.S., USSR).

Table 4: OLS regression estimating the relationship between dispute duration and the paper's main instruments. The results indicate that the instruments, on their own, exhibit little to no predictive power in explaining the duration of non-war interstate disputes.

Next, Table 5 reports pairwise correlations between dispute duration, the levels of hostility (ranging from no aggression to the use of force short of war), and the instruments. The correlations between the instruments and both duration and hostility levels are uniformly small in magnitude.

Taken together, these placebo results provide support for the exclusion restriction: the instruments are neither strongly correlated with dispute duration nor with proxies for the dispute's intensity. This reinforces the argument that their effect in the main specification operates through the negotiation environment specific to wartime dynamics.

	Host1	Host2	Host3	Duration	$r_{it}^{instig.}$	Instab$^{instig.}$	$r_{jt}^{instig.}$
Host1	1.000						
Host2	-0.089	1.000					
Host3	-0.173	-0.966	1.000				
Duration	-0.020	-0.109	0.113	1.000			
$r_{it}^{instig.}$	-0.047	0.177	-0.163	0.005	1.000		
Instab$^{instig.}$	-0.013	-0.058	0.061	0.059	-0.039	1.000	
$r_{jt}^{instig.}$	-0.013	0.016	-0.012	-0.041	0.180	-0.007	1.000
Instab^{target}	0.003	0.024	-0.024	-0.006	-0.020	0.068	-0.022

Notes: Pairwise correlations among variables used in the placebo test sample (non-war disputes after 1970). Host1, Host2, and Host3 represent levels of conflict intensity: no aggression, verbal threats only, and shows of force, respectively. $r_{it}^{instig.}$ and $r_{jt}^{instig.}$ refer to initial latent military power for the instigator and target, respectively. $Instab^{instig.}$ and $Instab^{target}$ denote initial social instability trends. Dispute duration is measured in months.

Table 5: Correlation between dispute duration, levels of hostility, and the paper's main instruments. The instruments exhibit little correlation with either dispute duration or dispute severity, consistent with the exclusion restriction.