Only Here to Help? Bargaining and the Perverse Incentives of International Institutions

**William Spaniel**

*University of Pittsburgh*

Many international organizations reduce the costs states incur from conflict. Critics argue that the expectation of such aid, by mitigating potential suffering, perversely incentivizes states to initiate conflict more often. I develop a model that formalizes this intuition. It shows that institutions may still ameliorate suffering in two ways. First, they may absorb so many of the expected costs that they compensate for the fact that conflicts occur more often. Second, aid can have a second-order effect of reducing uncertainty about the costs of conflict; to the extent that this uncertainty explains why parties cannot negotiate a cooperative alternative, aid might actually reduce the incentive to initiate conflict. Whether aid ultimately helps or hurts therefore depends on how it interacts with the informational challenges states face.

Consider the incentives of Syrian rebels on the eve of civil war. Some opponents of the Assad regime wanted to see democratic reforms and new policies to reduce economic inequality (Carpenter 2013; Coutts 2011). Violence provided one way to resolve the political dispute between the regime and its contentious citizens. Yet, most of the military power rested in the hands of the government. Thus, it seems reasonable to conclude that those about to rebel expected a costly conflict.

Nevertheless, rebel groups might also have anticipated that the international community would respond by supplying humanitarian assistance to civilians. Indeed, countless groups contributed after fighting began, including the United Nations (UN) High Commissioner for Refugees, the World Food Program, the World Health Organization, Doctors Without Borders, and the Red Cross. War remains deadly and destructive despite their efforts. Nevertheless, they have relieved some of the costs of war by providing food, shelter, and medical care.

Policymakers praise these cost-mitigating institutions. In a speech to the UN, then US Secretary of State John Kerry called the situation in Syria the "greatest humanitarian catastrophe since World War II" and underscored the importance of delivering aid without impediment (Walker 2016). Observers have commended these efforts more generally. The High Commissioner alone has twice won the Nobel Peace Prize.

However, skeptics warn of unintended consequences. In particular, many civil war scholars believe that conflicts result from bargaining failures and, therefore, reducing the expected costs of war can make turning to force a more attractive option (see Kuperman 2005, 2008; Rauchhaus 2009; Narang 2019). In turn, they may take more aggressive negotiating postures that precipitate conflict. This suggests that institutions and organizations that provide humanitarian aid perversely cause the exact problems they seek to solve.

A more general pattern extends beyond cases of war. Various institutions and organizations also provide assistance that affects the costs of arms development, trade disputes, and sanctions. Thus, we need to better understand the nature, and extent, of the perverse incentives at stake. In this article, I suggest some answers by formalizing the relevant dynamics. My results identify nuances overlooked in the existing literature. Critics correctly argue that aid provision may increase the likelihood of conflicts. As a consequence, the net costs suffered may also increase. Yet, champions of these efforts will also find reasons for optimism. Institutions that seek to mitigate the costs of conflict can also counterintuitively reduce conflict. Moreover, institutions may be so effective in their mitigation efforts that they reduce more costs than they cause.

To understand the incentives at work, I develop a model in which two actors bargain over a good. If they fail to reach an agreement, one party may capture that good. Both parties suffer costs as a result. An international institution observes these costs and exerts effort to reduce them. The more effort exerted, the less the parties ultimately suffer.

I begin with a case in which one party does not know the other’s cost of conflict. Institutional effort has asymmetric effects, so it may help to conceptualize the results using the Syrian civil war example. Suppose Syrian rebels did not know how much the government would concede. As the institution reduces the uncertain party’s (the rebels’) costs, the probability of conflict increases. This occurs because the institution protects that party from the dangers of bargaining failure. Correspondingly, that party assumes more risk by increasing its demands. The institution has the perverse effect that critics predict.

In contrast, mitigating the costs of the actor not facing uncertainty (the government) has an unexpected consequence. As a second-order effect, reducing costs in this manner also decreases the opponent’s (the rebels’) uncertainty. When the institution reduces that actor’s costs, the various types behave more similarly. This reduces the information here, the institution’s own preferences insensitive the additional risk. Regardless of how rebel actions led to civil war, for example, the United Nations High Commissioner for Refugees is willing to provide assistance to displaced individuals. Indeed, the model I develop has perfect information, whereas imperfect information is a necessary condition for moral hazard.

William Spaniel is an assistant professor of political science at the University of Pittsburgh.

**Author’s note:** I thank Phil Arena, Emiel Areal, Thad Brines, John Duggan, Louisa Eden, Douglas Farlows, James Frounge, Mark Fein, Hein Goeman, Matt Jackson, Gleeson Judd, Neil Narang, Michael Pozen, Ashwarya Ramesh, Scott Sagan, Burkett Savu, Brad Smith, Randall Stone, Emily Vanmeter, the editors of *International Studies Quarterly*, and the anonymous reviewers for their comments on previous drafts.

© The Author 2018. Published by Oxford University Press on behalf of the International Studies Association. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

---


---

This literature often describes this as moral hazard. However, the mechanisms are distinct (Rauchhaus 2009). In the classic setting, moral hazard occurs because an insurer cannot observe the policyholder’s actions and therefore cannot construct a complete contract. Embedded, the policyholder takes riskier actions.

---

Downloaded from https://academic.oup.com/isq/article-abstract/62/1/14/4797558 by Falk Library user on 01 June 2018
nuclear war and rogue nuclear attacks. The international community has addressed these on at least three fronts. First, International Atomic Energy Agency safeguards make theft of nuclear material more difficult. Second, the Proliferation Security Initiatives creates an international network to intercept nuclear traffickers (see Nikitin 2011). Finally, following the September 11 attacks, the UN Security Council passed Resolution 1540. This obligated countries to outlaw the possession and development of nonstate weapons of mass destruction (see Bosch and Van Ham 2007). The Group of Eight followed with the creation of the Global Partnership Against the Spread of Weapons and Materials of Mass Destruction (see Heyes, Bowen, and Chalmers 2012).

Aid provisions extend to issues of trade. Economic sanctions against North Korea limit Pyongyang’s ability to feed its population. The World Food Programme instead provides a portion of the country’s basic caloric intake. The organization took similar actions in the 1990s when sanctions against Serbia caused the same problem (Oshunrinade 2010, 21). Meanwhile, adjudication processes under the General Agreement on Tariffs and Trade and the World Trade Organization offer states a means to resolve trade disputes. Reaching an agreement requires compromise, but states may prefer that to suffering long periods of inefficiency that might result otherwise (Busch and Reinhardt 2000).

The various institutions I have discussed thus far address a variety of seemingly unrelated outcomes. Yet, they share a basic logic. Before institutions act, negotiations can ensure that costs from conflict never materialize. States, or governments and rebel groups, routinely bargain to avoid war (Fearon 1995); proliferators and their rivals can negotiate over policies to disincentivize nuclear development (Spaniel 2015; Bas and Coe 2016); the sender and the target of sanctions can compromise on an issue that might otherwise result in sanctions (Drezner 2000); trade partners can adjust their barriers to sustain trade flows (Reinhardt 2001). If these mutual adjustments fail to resolve the problem, a variety of institutions can exert effort to minimize the damage, but they can only do so once bargaining has failed—the High Commissioner cannot house war refugees without a war, for example.

How do the various strategic incentives interact? The qualitative and quantitative literatures, at least in regard to war, suggest an answer: not well. Broadly, researchers observe that cost mitigation decreases the pain parties suffer if negotiations fail. This perversely shrinks the range of mutually acceptable settlements, which seemingly implies that more conflict occurs as a consequence.

However, complex bargaining interactions often require “accounting standards” to sort out (Powell 1999, 29–34). To develop those standards, the substantive examples suggest a model with two distinct phases. For the reasons outlined above, the first phase must feature bilateral bargaining.

---

2 Note the distinction between these institutions and those that reduce the transaction costs for information gathering or distribution (Keohane 1984). Such costs do not stem from bargaining failure and do not mitigate harm. Although the timing of the institution’s reduction seems consequential, it proves irrelevant here. See Metrowitz, Morelli, Ramse, and Spinnato (2015) for a similar argument on how information provision causes additional conflict.

3 Broad institutions like the International Atomic Energy Agency (IAEA) do more than just mitigate costs, though I focus on this function because such institutions can choose whether to stunt those subdivisions. Like the High Commissioner, the IAEA has won a Nobel Peace Prize for its efforts. Critics of the inspections regime—including former director of Los Alamos National Laboratory Siegfried Hecker (2008)—argue that they place too much focus on government compliance with nuclear agreements and do not do enough to eliminate outright theft.

4 Unilaterally, the United States also offers to share its permissive action link technology. These tools ensure command and control for existing nuclear arsenals, reducing the chances of an accidental launch (Caldwell 1987).

5 This reflects broader declines in public health following economic sanctions (Gibbons and Garfield 1999; Allen and Leckie 2012).
The second phase should feature a fully-strategic institution that limits the damage if bargaining failed beforehand.6 This ordering prevents an institution from committing to an incredible level of assistance, in line with the qualitative literature's concern about perverse incentives. After establishing this baseline, I add asymmetric uncertainty to the model. Incomplete information allows for a more compelling setup and addresses a common cause of conflict (Schelling 1966, 99–105; Myerson and Satterthwaite 1983; Feanor 1995; Reinhardt 2001).

Developing the model in this manner highlights problems unique to cost mitigation. Previous research has explained how malfunctioning, misinformed, or poorly managed institutions can backfire (Luttak 1999; Terry 2002; Valentin 2011; Hawkins, Lake, Nielsen, and Tierney 2006). In contrast, the institutions I analyze function as intended: if inefficiency exists, any effort the institution exerts mitigates the damage. Focusing on this best-case scenario stacks the deck in favor of institutions improving outcomes. Pessimists will find greater vindication if institutions indeed backfire under these relatively rosier conditions.

A related literature argues that shifting the distributional gains of conflict primes institutions for failure (Regan 2000; Anderson 2010; Wood and Molfino 2016). In sum, scholars worry that such institutions incentivize the beneficiary to misbehave. For example, Lisher (2005) argues that some refugee camps provide sanctuary to rebel militaries, allowing them to regroup and increase their chances of defeating the government. Reporting that civil wars featuring interventions extend the length of conflict, Regan (2009) finds broader empirical support for a similar mechanism. Elbadawi (1999) further argues that anticipated interventions allow weak rebel groups to initiate otherwise unwinnable conflicts.

One may wonder whether those results persist with an institution with no preference for the distributive outcome.7 This would again seem to stack the deck against perverse effects. Developing the model accordingly allows me to contrast my work from two related models. In Chapman and Wolford (2010), a challenger can consult a biased institution, which can hinder the challenger's mobilization efforts. It similarly contrasts with Kydd and Straus’s (2013) work, in which an intervener has a moderate preference over the outcome and can enter the war to impose its ideal point. These models find that the third party sometimes creates perverse incentives. I find that institutions can cause conflict even without altering the expected outcome (Cunningham 2016).

Uncertainty over the Costs of Conflict

The game consists of three actors: two states and an institution.8 Broadly, the states negotiate over control of a disputed good, while the institution exerts effort to reduce burdens should bargaining break down. Play begins with state 1 demanding $x \in [0, 1]$ of the good, which I standardize to value 1 without loss of generality. State 2 sees the demand and accepts or rejects it. Accepting ends the game with state 1 receiving a $x$ and state 2 receiving the remaining $1-x$.

If state 2 rejects, it pursues an inefficient policy to capture the good for itself. This transfers the pie to state 2 but creates costs $c_1, c_2 > 0$ for the players. At this point, the institution exerts effort to minimize those costs. Specifically, it chooses levels $\alpha_1 \in \{0, 1\}$ and $\alpha_2 \in \{0, 1\}$ to reduce state 1’s cost burden to $\alpha_2 c_2$. Thus, state 1’s overall payoff for bargaining breakdown equals $\alpha_1 c_1$, while state 2’s equals $1 - \alpha_2 c_2$. To focus on the interesting cases (that is, situations in which state 2 has a credible threat to reject), I restrict attention to parameters in which state 2 has a positive utility for rejecting even if the institution takes no action.

I impose minimal structure on the institution’s payoff function to keep the results general. The institution has a tradeoff between minimizing inefficiency and exerting effort of its own. Let $g(\alpha_i)$ represent how much the institution internalizes state 1’s cost, where $\frac{d}{d\alpha_i} g > 0$ and $\frac{d^2}{d\alpha_i^2} g \geq 0$. Substantially, this means that the larger the share of costs state 1 actually suffers (that is, as $\alpha_i$ increases), the more the institution internalizes that state’s cost. It also means that the institution internalizes the extra burden of a given percent at least as much as it internalizes the next percent.

The institution faces an effort constraint, however. Let $h(\alpha_i)$ represent the institution’s burden from exerting some amount of effort to reduce state 1’s cost, where $\frac{d}{d\alpha_i} h < 0$ and $\frac{d^2}{d\alpha_i^2} h \geq 0$. Substantially, this means that the institution finds less effort (i.e., $\alpha_i$ close to 1) cheaper than more effort and that each additional percent becomes increasingly costly to reduce.9 Additionally, I require that at least one of the constraints on the second derivatives of $g$ and $h$ holds strictly.

All told, the institution’s utility equals $-g(\alpha_1) + g(\alpha_2) - h(\alpha_1) - h(\alpha_2)$. Given this utility structure, the institution may prefer reducing one side’s costs by a greater percentage than the other’s.

Complete Information Equilibrium

Going through the baseline complete information case generates some intuition for what follows. Backward induction yields the solution. The institution moves last and has complete and perfect information. It therefore faces a constrained optimization problem on its selections for $\alpha_1$ and $\alpha_2$. Despite the general form of the institution’s utility function, a unique solution pair exists, which I call $\bar{\alpha}_1$ and $\bar{\alpha}_2$.

If the effort function does not cause too much of a loss, these optimal mitigation strategies fall strictly below 1. That is, the institution reduces the states’ costs provided that it has a minimal capacity to do so.

From here, the states treat their costs of bargaining breakdown as $\bar{\alpha}^2 c_1$ and $\bar{\alpha}^2 c_2$ because they actually suffer those amounts if state 2 rejects.10 Because bargaining implies inefficiency, a range of mutually satisfactory demands exists. Thus, in the ultimatum game setup, state 1 demands the most state 2 accepts, which equals $\bar{\alpha}^2 c_2$; state 2 accepts.

Consequently, the institution’s move has no effect on whether bargaining breaks down with complete information. Instead, it merely alters the terms of the settlement.

---

6 A fully strategic institution ensures that the perverse effects follow from optimal play and not a behavioral assumption built into the model.

7 Whereas previous work mixes the two effects (Koperman 2008), I isolate the effect of cost mitigation by considering it in isolation. See Cunningham (2016) for a counterexample as to how potential interventions decrease civil war onset.

8 For civil war cases, one might rebuff a state as a rebel group.

9 For example, after having paid the fixed costs to create organizational infrastructure, institutions can more easily create the first refugee camp than the second. This may result from international institutions picking more hospitable and less dangerous locations for their initial efforts.

10 Analogously, $\bar{\alpha}^2$ and $\bar{\alpha}^2$ represent the expected portion of costs in a game where both state 1 and state 2 face manual uncertainty over the institution’s preferences. This holds because the institution moves last, so no signaling dynamics exist as long as the states face equal uncertainty. The states can only maximize over what they expect to occur, which generates an analogous proof to the complete information case.
Incomplete Information Equilibrium

States may not know how costly their opponents view outside options. To incorporate this into the model, suppose state 1 faces uncertainty about state 2’s cost $c_2$. Specifically, nature begins the game by drawing $c_2$ from a continuous and strictly increasing cumulative distribution function $F(c_2)$ on the interval $[c_{1l}, c_{2u}]$. $F(c_2)$ is differentiable everywhere on the interval and has density $f(c_2)$. The results below apply to distributions with strictly increasing hazard rates that create sufficiently responsive optimal demands. The uniform distribution falls in this class. State 2 observes the draw, but state 1 only knows the prior.

Moving forward, suppose that even if the institution does not observe the draw, it still optimizes according to state 2’s realized cost. This assumption has two justifications. First, state 1’s uncertainty may stem from not knowing state 2’s “resolve”—that is, how much state 2 values the good. Due to the standardization of the good at 1, the model incorporates the value into the cost state 2 pays. Thus, state 1, state 2, and the institution may all have common knowledge of the actual cost of bargaining breakdown, just not how much state 2 values it relative to the good at state. In turn, the institution can observe the inefficiency and reduce it. Second, the institution can see problems as they happen and address them. For example, imagine state 1 faced uncertainty about whether state 2 would suffer a massive humanitarian crisis after bargaining breakdown. The institution may not know this during the bargaining phase, but it can observe and address the crisis as conflict continues.

Given that the institution still has a straightforward optimization problem, it selects $\alpha_1^e$ and $\alpha_2^e$. Although $\alpha_2^e$ interacts with the source of state 1’s uncertainty ($c_2$), depending on the other parameters, the information asymmetry can induce state 1 to make a demand that causes bargaining breakdown:

**Proposition 1:** If $\alpha_1^e > \alpha_2^e$, the states reach an agreement with certainty. If $\alpha_1^e = \alpha_2^e$, state 1 offers an amount that state 2 rejects with positive probability. Inefficiency results in this case.

The appendix contains a complete proof. Scholars familiar with bargaining models featuring uncertainty will recognize this as a classic risk-return tradeoff. The more state 1 demands, the more it keeps conditional on state 2 accepting. But greater demands simultaneously make state 2 more likely to reject, forcing state 1 to pay costs. Consequently, state 1 must balance between extracting the best deal possible and avoiding bargaining breakdown.

Intuitively, when bargaining breakdown appears relatively more painful, state 1 makes a safer demand. Thus, if $\alpha_1^e$ is sufficiently great, it demands $\alpha_2^e$. This induces the lowest cost type of state 2 to accept, which also ensures that all other types accept. For example, state 1 will not risk a breakdown to capture a slightly larger share of territory if war will destroy the homeland. But at smaller $\alpha_1$ values, state 1 gambles on a demand that state 2 might reject.

**Comparative Statics and Empirical Implications**

The intuition for proposition 1 appears to vindicate scholars who argue that cost mitigation incentivizes bargaining breakdown. The first comparative static provides formal validation:

**Proposition 2:** As the institution reduces state 1’s costs by a greater portion (that is, as $\alpha_1^e$ decreases), the probability of bargaining breakdown weakly increases.

The explanation for proposition 1 gives the intuition. Decreasing the portion of costs that state 1 internalizes (that is, decreasing $\alpha_1^e$) reduces state 1’s punishment for bargaining breakdown. Insulated, state 1 runs greater risks. For example, imagine that state 1 in the absence of institutional cooperation would have demanded an amount state 2 certainly would have accepted. Slight institutional help (that is, picking $\alpha_1^e$ close to 1) fails to change this. But greater assistance convinces state 1 to increase its demand and run some risk. No bargaining breakdown occurred before. Now, rejection—and the corresponding inefficiency—occurs with positive probability.

A similar problem extends to situations where state 1 would have risked rejection in the absence of the institution. Here, the assistance causes state 1 to further increase its demands. Thus, some types of state 2 that would have found state 1’s original demand acceptable find the new demand unacceptable. These types add to the overall probability of bargaining breakdown. The effect compounds as $\alpha_1^e$ moves closer to 0.

Civil war scholars have flagged this mechanism and understand the basic intuition. If a rebel group knows that the international community will feed and shelter its civilians, war looks more attractive. The group responds by increasing its demands on the government. Doing so implies a greater risk of bargaining breakdown, but the institution has made that risk acceptable.

Even so, a crosscutting effect exists. The institution promotes inefficiency by causing state 1 to issue more aggressive demands, leading both states to pay costs more frequently. Yet, the institution also promotes efficiency by reducing the portion of costs ultimately suffered. One may think the former effect dominates, but the following model qualifies that intuition:

**Remark 1:** The expected costs paid are nonmonotonic in $\alpha_1^e$. In particular, initial reductions in state 1’s cost can increase overall inefficiency, but larger reductions can decrease inefficiency.

Figure 1 illustrates the intuition. It looks at the overall expected costs paid ($\alpha_1^e + \alpha_2^e$) as a function of $c_1$ for two cases. In one case, $\alpha_1^e = \frac{1}{2}$ and $\alpha_2^e = \frac{1}{2}$. To simulate an environment without the institution, $\alpha_1^e = \alpha_2^e = 1$ in the other case. Let state 2’s cost be uniformly distributed on the interval $[\frac{1}{2}, \frac{3}{2}]$. In each case, state 1 demands an amount that state 2 certainly accepts if state 1 faces a sufficiently large cost. Without institutions, that critical cost level is $c_1 = \frac{1}{3}$.

---

11 A careful reader will note that the rest of the model essentially becomes an exercise in comparative statics on $\alpha_1^e$ and $\alpha_2^e$; an interesting result in its own right.

12 Formally, I assumed the hazard rate $\frac{f(c_2)}{1-F(c_2)}$ strictly increasing and that $F(c_2)$ is differentiable everywhere on the interval $[c_{1l}, c_{2u}]$, where $F(c_2)$ defines the implicit function that maps $c_2$ to state 1’s optimal demand. Later, I show that a limiting result applies to distributions without these properties.

13 Consequently, the relationship in proposition 2 is weakly increasing.
with institutions, it is $c_1 = \frac{1}{1 + \gamma}$. This smaller amount reflects proposition 2's claim that absorbing state 1's cost increases the probability of bargaining breakdown.

Below those critical cost thresholds, state 1 allows for positive probability of rejection. For values of $c_1$ between $\frac{1}{2}$ and $\frac{1}{3}$, an efficiency gap appears because bargaining breakdown only occurs with the institution. Thus, though the institution mitigates some of the damage, it backfires. This principle continues for some $c_1$ values lower than $\frac{1}{3}$, as the institution causes more bargaining breakdown than it can compensate for.

Why does the institution behave in this manner if doing so hurts its welfare? Some organizations pay little attention to the politics surrounding a conflict (Lischer 2005, 4). But the model highlights a second mechanism: institutions face a credibility problem. At any point during the game, the institution wants to minimize costs and externalities. This creates a time-consistency preferences issue. In the institution's ideal world, it would commit to not state 1's costs to eliminate the perverse incentive. However, credibility issues doom this solution—if the institution has a preference to reduce inefficiencies, it must exert effort in the event of bargaining breakdown.14 Or, in the words of a High Commissioner for Refugees official, "even the guilty need to be fed" (quoted in Rief 2002, 54).15

The credibility problem has upstream consequences. Recognizing the incredible commitment, state 1 knows the institution will offer assistance. In turn, state 1 increases its demand, knowing that the institution will mitigate some of the damage if bargaining fails.16 It just so happened that the preference to reduce costs inadvertently creates them.17

**Figure 1.** The expected inefficiency with and without the institution as a function of State 1’s cost $c_1$

**Figure 2.** The difference in efficiency with and without the institution

Fortunately, something interesting happens as $c_1$ approaches 0. At this point, state 1 pays few costs for bargaining breakdown in the absence of the institution. It therefore demands an amount that many types reject. Introducing the institution causes state 1 to expand it demands, but only slightly—with costs small already, the institution cannot greatly impact state 1's negotiation strategy. More types now reject. Nevertheless, the portion of rejections attributable to the institution pales in comparison to the portion of rejections state 1 instigates on its own.

Meanwhile, in the absence of the institution, state 2 pays its full costs. But with the institution, state 2 only pays a fraction. Because the institution reduces the cost of the types that would have rejected anyway, the institution improves outcomes. That is, the institution promotes efficiency despite perversely causing bargaining breakdown. Figure 2 illustrates this by looking at the net difference between a world with the institution and a world without.

The existing literature often overlooks this net-positive effect. Yet, the model indicates that demonstrating that an institution caused bargaining breakdown does not prove that the institution backfired. Rather, the institution may have more than compensated for its deficiency by fixing problems that were likely without it (Western 2005). Thus, a proper review of an institution requires investigating the bargaining and implementation phases in both the actual and counterfactual world.

---

14This distinguishes cost mitigation from mechanisms that require disputants to agree to them ex ante, like the laws of war (Morrow 2014). A repeated interaction could solve the institution's problem, but this requires a high probability of conflict in any given period. Low probability events like many of the conflicts discussed here fail to qualify.

15Institutional chimeras commonly lack this conditionality. For example, the Mine Action Service's mission statement seeks a "world free of the threat of landmines and unexploded ordinance"; nowhere does it suggest that it would refuse calls for assistance.

16This contrasts with information-oriented institutions as in Chapman and Wolford (2010). These institutions alter the costs of conflict by offering support during the bargaining process. Such timing eliminates the credibility problem and allows a fully strategic institution to only induce more conflict if it wants to.

17This connects to the broader literature on how institutions alter the terms of agreements or lead to unintended consequences (Lischer 2005; Yuen 2009; Murdie and Davis 2010; Benson, Meloniou, and Runyan 2014; Johns and Pfei 2014).
Figure 2 also raises more general questions about when institutions improve outcomes. Scholars recognize that reducing costs and externalities can create benefits (Oye 1992; Mitchell and Kelibach 2001). Milner (1997, 44), for example, makes the seemingly straightforward claim that "[i]f ... externalities rise, ceteris paribus, so do the gains from cooperation, and hence the incentives for it." But my model counterintuitively shows the opposite: the institution can only reduce efficiency when state 1 has great costs; the potential for great inefficiency compels state 1 to work harder to avoid it. Gains instead arise when state 1 suffers small costs of conflict, which induces state 1 to issue more reckless demands.

Switching gears, this analysis so far has only investigated how changing the institution’s effort for the uninformed proposer alters efficiency dynamics. One might anticipate an identical problem as the institution mitigates more of the informed state’s costs. Yet, the following proposition brings great news:

**Proposition 3:** As the institution reduces state 1’s costs by a greater portion (that is, as \( \alpha_2 \) decreases), the probability of bargaining breakdown weakly decreases.

None of the perverse effects seen in proposition 2 arise here; the more the institution works to reduce state 2’s cost, the less often bargaining breakdown occurs.\(^{15}\)

The literature skeptical of international institutions has not yet identified these positive effects. Narang (2015), for example, theorizes that reducing the costs of war can only result in a higher probability of inefficient behavior. Nevertheless, proposition 3’s mechanism demonstrates the utility of formalization. Decreasing \( \alpha_2 \) reduces the cost that state 2 internalizes. All else equal, state 1 now finds reaching an agreement less attractive, as it cannot extract as much surplus. The existing literature predicts that state 1 would pursue a more aggressive bargaining strategy, leading to more bargaining breakdown and inefficiency.

However, all else is not equal. The proposer does not know state 2’s cost in this model. Strategically, uncertainty only matters in how it affects a player’s behavior. The unknown costs impact state 1 because it cannot anticipate which demands state 2 rejects. Manipulating \( \alpha_2 \) affects this and leads to the unexpected improvement.

To see why, consider the range of maximally acceptable demands with and without the institution. Without the institution, the lowest cost type accepts up to \( \epsilon_2 \), while the highest cost type accepts up to \( \delta_2 \). The magnitude of the discrepancy equals \( \delta_2 - \epsilon_2 \). With the institution, the lowest cost type accepts up to \( \alpha_2 \epsilon_2 \), while the highest cost type accepts up to \( \alpha_2 \delta_2 \). The magnitude of this new range of reservation values equals \( \alpha_2 \delta_2 - \alpha_2 \epsilon_2 \).

If \( \alpha_2 < 1 \), the institution reduces this range. This shrinks the variance of state 2’s possible reservation values.\(^{16}\) Put differently, the institution alleviates state 1’s information problem.\(^{17}\) Because the institution shrinks the difference between safe and aggressive demands, gambling looks less attractive. In turn, state 1 demands an amount that state 2 accepts more often.\(^{21}\)

This mechanism distinguishes itself from the informational argument common to the broader literature on international institutions. Standard liberal theory argues that institutions increase information by improving the quality of signals or by reducing the transaction costs of information acquisition (Keohane 1984; Koremenos, Lipson, and Snidal 2001; Chapman and Reiter 2004; Fang 2008). Here, however, the institution provides no information; with or without it, state 1 still has the same belief given by the distribution function. But the institution nevertheless helps because it shrinks the problem that incomplete information poses.

Pressing forward, the good news continues. The institution not only reduces the probability of bargaining breakdown, it also mitigates some of the problems in the less likely scenario that the states enter conflict, as explained in the following remark:

**Remark 2:** As \( \alpha_2 \) decreases, the expected costs paid are weakly decreasing.

Because the institution does not cause additional bargaining breakdown as \( \alpha_2 \) decreases, it cannot increase inefficiencies. On the contrary, when bargaining fails, the institution meets its goals by preserving some of the pie.\(^{22}\)

The differences between propositions 2 and 3 underscore the need to better understand the correlates of uncertainty in international processes. Institutions can reduce bargaining breakdown when they make various types behave similarly. Thus, predicting whether an institution will positively or negatively affect a bargaining process requires first knowing where the uncertainty lies.

Developments in the civil war literature indicate a problem here. The previous section detailed various efforts to support refugees fleeing from rebel or insurgent areas. A separate literature argues that rebel groups face uncertainty about their governments’ costs to maintain territories, especially against multiple potential challengers (Walter 2006, 2009a). Under these conditions, the institution disproportionately assists the uncertain actor. The model indicates that institutions cause problems here.

Placing additional structure on the institution’s utility function furthers this issue. The original setup permitted an independent optimization problem for each state’s cost. One may alternatively conceptualize the institution’s effort levels as interdependently constrained. For example, constructing a refugee camp for or providing humanitarian aid to a separatist population could decrease services to the government.\(^{23}\) One may then wonder how the results presented here would change under these circumstances.

In the appendix, I put some structure on this question. To summarize, suppose an institution cares more about helping a rebel group than a government. Intuitively, the institution variance drives the differences between propositions 2 and 3, not the fact that state 1 plays the proposer and state 2 plays the receiver in this particular model. See Reed (2003) for a similar discussion of variance and uncertainty.

\(^{15}\) Considering the extreme case may prove helpful. If \( \alpha_2 = 0 \), the variance collapses to 0, and state 1 faces no effective uncertainty whatsoever. It knows that 0 represents the maximal demand that all types of state 2 accept. State 1 faces no risk-return tradeoff here; it demands 0, which results in guaranteed acceptance. No inefficiencies occur. Note that this limiting result applies regardless of \( f(x) \)’s functional form.

\(^{16}\) The relationship is weakly increasing because the institution has no effect if the states would strike a bargain with certainty regardless.

\(^{21}\) These problems can result from powerful countries with disproportionate control over international institutions (Stone 2011) that use power to provide greater aid to their friends.
spends more to help the rebels and less to help the govern-
ment as that bias increases; greater concern for the rebels
makes exerting marginal effort on them more attractive. But
this extra effort must come from somewhere, and so the in-
stitution decreases its effort toward the government.

Unfortunately, this replacement effect creates a double
whammy for the informational structure the civil war lit-
erature emphasizes. Per proposition 2, alleviating the pain
of the uninformed party (state 1 in the model, the rebels
here) increases the probability of bargaining breakdown.
Per proposition 3, alleviating the pain of the informed party
(state 2 in the model, the government here) decreases the
probability of bargaining breakdown. So as an institution’s
inclination to help the rebels increases, the institution ex-
acerbates proposition 2’s problem while reducing the help
from proposition 3. This validates the literature’s concern
about unintended effects of civil war meddling.

Uncertainty over the Outcome of Conflict

Scholars also recognize outcome uncertainty as a major
cause of conflict. This appears most prominently in re-
search on warfare, when states do not know their oppo-
ents’ military capabilities (Fearon 1995; Slaughter 2003;
Powell 2004).\footnote{This may prove the more critical case—
whereas costly signals can fix uncertainty about resolve
(Fearon 1997; Slaughter 2011), they fail to do so for disrup-
tional uncertainty (Arena 2013).} Such uncertainty extends to
other contexts as well. For example, a state that implements sanctions
may not know the expected resolution due to uncertainty about
the opposing state’s capacity to weather the sanctions or the
opposing leadership’s stability (Eaton and Kugler 1999; Spaniel
and Smith 2015). Likewise, states may not know the
outcome of a trade case in front of international adjudica-
tion or whether their opponents will comply (Reinhardt
2001).

To analyze such a scenario, consider the following revised
model. The overall structure remains identical. State 1 de-
mands \( x \in [0, 1] \). State 2 accepts or rejects. Rejecting costs
creates costs \( a_1, a_2 > 0 \) for both players. Following rejection, an
institution (with the same payoff function as before) mitigates
the costs by choosing \( a_1, a_2 \in [0, 1] \). The states only pay
\( a_1x \) and \( a_2x \).

Unlike before, the states have common knowledge of the
costs. Instead, nature draws \( p \in \{ p_1, p_2 \} \), which represents
the portion of the good state 1 expects to receive in the
event of bargaining breakdown. State 2 receives \( 1 - p \). The
bargaining model of war literature refers to this as state 1’s
probability of victory, but one may equivalently think of it
as the expected outcome of the inefficient process.\footnote{This
could include a stalemate in war or the persistence of the status
quo following sanctions.} As before, state 2 observes the
draw, but state 1 does not. Let \( F(p) \) represent the continuous
and strictly increasing cumulative distribution of \( p \), and let \( f(p) \) be the corresponding
density function. I make three additional assumptions
about the distribution: it has full support on the interval
\( [p_1, p_2] \), is differentiable everywhere on the interval, and
has a strictly increasing hazard rate.\footnote{Formally, \( \frac{f(p)}{1 - F(p)} \) is strictly increasing in \( p \).}

This generates a key result:

**Proposition 4:** With uncertainty over the distribution of the good
via conflict, the probability of bargaining breakdown increases as
\( a_1^* \) or \( a_2^* \) decreases.

Put differently, institutional effectiveness can only in-
crease the probability of conflict. The previous findings
give the intuition. The probability of bargaining breakdown
is decreased under proposition 3 because of a second-order
effect — reducing costs shrinks the proposer’s uncertainty by
collapsing the type distribution. Reducing costs has no analog-
ous effect with outcome uncertainty. As such, reducing ei-
ther side’s cost increases the probability of bargaining
breakdown; decreasing \( a_1^* \) insulates state 1 from rejection, while
decreasing \( a_2^* \) makes extracting the entire (smaller) surplus
with certainty look less attractive.

This result lacks the nuance from the results on cost un-
certainty. Correspondingly, one may wonder whether they
extend to other extensive forms with outcome uncertainty.\footnote{See
Leventoglu and Tarar (2008) for an example of how different crisis
bargaining protocols yield different equilibrium predictions.}

I address this question in the appendix using Bayesian
mechanism design, which shows what must hold for equilib-
ria across a range of conflict games (Fey and Ramsay 2011). A
negative finding results under the most idyllic of circum-
cstances: if a bargaining outcome guaranteed efficiency (be-
cause the states reach an agreement with certainty), suffi-
ciently strong institutions (that is, institutions that pro-
duce sufficiently small values for \( a_1^* \) or \( a_2^* \)) cause bargaining
breakdown and inefficiency with positive probability.

Once again, institutions designed to mitigate the costs
of civil wars seem to cause problems. Beyond uncertainty
over the government’s cost of war, the literature points to
asymmetric information about a rebel group’s capabilities as
the other uncertainty-based explanation.\footnote{The covert nature of
many rebel groups masks its true capabilities (Walter 1999, 132;
Fearon 2007; Walter 2003, 248–50) and obscures the expected
outcome of war.} The model unambiguously shows that institutions
cause bargaining breakdown in such environments.

Nevertheless, as seen earlier, additional bargaining failure
does not imply greater expected inefficiencies. The follow-
ing remark describes the silver lining:

**Remark 3:** With uncertainty over the distribution of the good via
conflict, the expected costs paid are nonmonotonic in \( a_1^* \) and \( a_2^* \).

In particular, initial reductions to either state’s cost can increase
overall inefficiencies, though larger reductions can eventually lead
to an overall decrease.

As before, although decreasing costs perversely causes
more bargaining breakdown, powerful institutions see their
savings cover the problems they cause. Unfortunately, the
benefits accrue slowly here—without a second-order reduc-
tion of uncertainty, cost savings can only come as a direct
result of the institution’s effort.

Conclusion

This article explored whether international institutions en-
haracterize giving the costs of conflict. Previous work suggests
they may only incentivize bargainers to behave more recklessly in nego-
tiations, thereby preventing conflicts that would not occur in the
institution’s absence. I developed a formal model to scruti-
nize the claim. Although the model finds the anticipated ef-
fect, it also shows that the institution can enhance efficiency
in two ways: (1) it can resolve information problems by re-
ducing the variance of potential reservation values and (2)
it may absorb so much of the expected costs that it compen-
sates for the additional inefficiency it causes.

---

Downloaded from https://academic.oup.com/ijig/article-abstract/62/1/14/4797558
by Fethi Library user on 01 June 2018
These caveats matter for those involved in managing and designing efforts to reduce the negative costs of conflict. Humanitarian aid can increase bargaining failure when the expected costs of conflict are high; such harm-mitigating institutions can only overcome their own inefficiencies when they help an actor that already expects relatively low costs from conflict. Yet the same institutions tend to target the biggest crises—that is, where the existing or potential suffering from conflict is greatest. My model indicates that some of this effort is misguided. Ironically, such institutions prove most effective when they absorb low-level inefficiencies. Adhering to this recommendation may prove difficult, though, due to lack of political will among donor countries for off-the-radar issues.

This problem also raises questions about institutional design. Although some cost-mitigating institutions do not require substantial investment before any conflict, others do. Thus, if institutions can cause bilateral inefficiencies, why develop them in the first place? The "rational design" argument (Koremenos, Lipson, and Snidal 2001) suggests an answer: policymakers promote institutions because the positive of doing so outweighs the negative. This could be for a number of reasons. In the aggregate, institutions may stop more conflicts than they cause. Alternatively, even if they cause more conflicts, they may absorb more damage in expectation than they add. Building the institution can also fix preexisting problems at the cost of causing future problems. In contrast, a pessimistic view argues that powerful states capture these institutions (Stone 2011). This could cause those powerful states or their allies to benefit from the mitigation despite the additional instability. Still, future work could investigate whether states can temper the drawbacks, bearing in mind the constraint of the credibility problem.

Finally, an interesting avenue for future research would be to investigate similar institutions that manipulate conflict by imposing costs. For example, some postconflict peace missions maintain stability by militarily punishing violators (Doyle and Sambanis 2000; Fortna 2004; Mattes and Savun 2009). However, my results indicate that mitigation does not always have a pacifying effect. Threatening violators with punishment might alter bargaining dynamics and cause the side facing uncertainty to take greater chances. Fortunately, these institutions have the benefit of manipulating ongoing negotiations. But my model suggests that finding the right incentive structures my prove difficult.

Supplemental Information

Supplementary information is available at williamspaniel.com and at the International Studies Quarterly data archive.

References


