Uncertainty in Crisis Bargaining with Multiple Policy Options*

Brenton Kenkel† Peter Schram‡

February 7, 2022

Abstract

We examine the relationship between private capabilities and war in a new class of crisis bargaining games. Whereas traditional models consider interactions that end in either war or a peaceful bargain, we assume actors can also engage in low-level, costly policy options that shape final political outcomes in their favor (e.g., sanctions, airstrikes, cyberattacks, etc). We analyze these *flexible-response crisis bargaining games* using the tools of mechanism design. In contrast with standard monotonicity results in crisis bargaining models, we identify general conditions under which improving a state's private willingness to fight may decrease the probability of war or that state's expected utility from a settlement. Theoretically, our results identify when including multiple conflict options breaks key tenants regarding the role of private information in conflict. Substantively, our results suggest that improving specific conflict capabilities can perversely lead to more war or worse outcomes for the state making the improvements.

*We thank Andrew Coe, Brad Smith, William Spaniel, and participants at the Vanderbilt Conflict Workshop for helpful comments.

†Assistant Professor, Department of Political Science, Vanderbilt University. `brenton.kenkel@vanderbilt.edu`

‡Assistant Professor, Department of Political Science, Vanderbilt University. `peter.schram@vanderbilt.edu`

"Our traditional approach is either we're at peace or at conflict. And I think that's insufficient to deal with the actors that actually seek to advance their interests while avoiding our strengths."
—General Dunford (2016), Chairman of the U.S. Joint Chiefs of Staff.

The relationship between private information and war is central to contemporary international relations theory (Fearon, 1995; Powell, 1999; Schultz, 1999; Wagner, 2000; Slantchev, 2003b; Meirowitz *et al.*, 2008; Chassang and Padró i Miquel, 2009; Acharya *et al.*, 2015; Ramsay, 2017; Gurantz and Hirsch, 2017; Spaniel and Bils, 2018). Theories of how private information affects war are typically developed using crisis bargaining games, in which states engage in a series of negotiations over a divisible asset and each state may go to war as an outside option.[1] These models assume that one or more states receive a private signal about their war payoffs, unknown to their adversaries, that represents their hidden capabilities or willingness to go to war. Across the diverse set of models of crisis bargaining, two monotonicity results consistently emerge. First, when an actor receives a better private signal about its war payoffs, negotiations are more likely to end in war. Second, when an actor receives a better private signal about its war payoffs, that actor will attain a greater final expected utility. In two seminal papers, Banks (1990) and Fey and Ramsay (2011) utilize the tools of mechanism design to assess just how ubiquitous these monotonicity results are. They find that the monotonic relationships between private type and outcomes are not specific to any particular bargaining game form or equilibria, but instead emerge in *every* equilibrium of *any* crisis bargaining game.[2]

The crisis bargaining framework has shaped the theoretical study of the causes of war, but traditional models rest on an implausible assumption: that actors either reach a peaceful and efficient bargain, or else engage in full-scale war (or some similar inefficient engagement). Recent research has relaxed this dichotomy, recognizing that crisis actors in practice face a vast array of policy options between completely efficient peace and totally destructive war. These policy options include implementing sanctions or tariffs (Coe, 2014; McCormack and Pascoe, 2017; Spaniel and Malone, 2019; Joseph, 2020); offering third-party support to rebels (Schultz, 2010); pursuing a wide range of low-level operations sometimes classified as gray zone conflict (Mazarr, 2015; Gannon *et al.*, 2020), hybrid conflict (Lanoszka, 2016) or hassling (**?**Schram, 2021); engaging in cyberwarfare (Gartzke and Lindsay, 2015; Baliga *et al.*, 2020); entering into brinkmanship (Powell, 1989, 2015); and arming (Schultz, 2010;

[1] We use the broad definition of crisis bargaining games from Fey and Ramsay (2011).

[2] The monotonicity results do not hold for models—for example Slantchev (2011)—which considers an alternate form of uncertaninty than Fey and Ramsay (2011).

Debs and Monteiro, 2014; Gurantz and Hirsch, 2017; Coe and Vaynman, 2020).

Once we allow for flexible policy options in between total peace and total war, the concept of a private willingness to go to war becomes more complicated. In standard crisis bargaining models, private information is unidimensional, only affecting the payoffs from total war. However, what makes a state better at war might also affect its success using these intermediate policy options. For example, if a state has a wide range of privately known cyber-exploits, then the state knows that it could perform well in a conventional war that uses cyberattacks or in a precise cyberattack against a target's infrastructure. In other words, when a state possesses strong private cyber-capabilities, both war and low-level conflict could be better options for the state. Alternately, if a leader is privately concerned about losing popular domestic support, then the leader may be more willing to fight a war to create a rally-around-the-flag effect rather than implement tariffs or a covert low-level attack (Baker and Oneal, 2001; Chapman and Reiter, 2004). In other words, having a strong private desire to garner domestic support could make war a better option while making low-level conflict or economic warfare worse options. Whether private war payoffs are associated with greater or lower payoffs from alternative policy options is an empirical question—one whose answer varies across cases and contexts. What is important for our purposes is that these linkages undoubtedly exist, and their effects on the outcomes of crisis bargaining have not been systematically examined.

To summarize, contemporary international relations research has begun embracing that actors face a wide range of potentially-related policy choices within a crisis. But previous attempts to draw general conclusions about crisis bargaining models have implicitly ruled out this diverse array of policy choices (Banks, 1990; Fey and Ramsay, 2011). We bring these literatures together.

This paper re-examines the relationship between private information and war within what we call *flexible-response crisis bargaining games*. These games share the key features of crisis bargaining games, but also allow for multiple forms of costly conflict short of all-out war. Importantly, we also allow the private type to influence the payoffs of both war and of the other forms of conflict. We find that previously established monotonicity results for crisis bargaining models, like the relationship between greater private war payoffs and a greater final likelihood of war, break down in the flexible-response models. Using a game-free analysis along the lines of previous mechanism design research (Banks, 1990; Fey and Ramsay, 2011), we show that these differences with conventional findings are not the result of any particular game form or functional form. Instead, these differences can arise in *any* flexible-response crisis bargaining game, depending on particular assumptions about modeling primitives and the strategies available to players short of war.

Flexible-response crisis bargaining games consider two competing players, a challenger and a defender, in a crisis. These games begins with the challenger selecting some level of transgressions (possibly none) where, following Gurantz and Hirsch's (2017) use of the term, the transgression is some act that is beneficial to the challenger but hurts the defender.[3] In response, the defender can enter into a decisive war over the transgression, can allow the transgression to come to fruition, or can "hassle." Hassling, following **?**'s (**?**) use of the term, undercuts the transgression through some destructive and non-decisive low-level response. How the defender plays the game is a function of their private type, which influences their war and hassling payoffs. Beyond this, we place no particular structure on the game. Within flexible-response crisis bargaining games, states could bargain, send costly or non-costly signals, make ultimatum offers, walk back or increase their selected hassling levels, or some combination of any of these actions before the game eventually ends.

As a "proof of concept" to illustrate how adding additional conflict options can undermine the canonical monotonicity results, consider the models in Figures 1 and 2. In both games, Nature moves first and designates D as type $\underline{\theta}$ with probability $\Pr(\underline{\theta}) \in (0,1)$ and as type $\bar{\theta}$ with probability $1 - \Pr(\underline{\theta})$. D knows their own type, but C does not. The higher type of D has a higher war payoff. Next, C selects whether to transgress ($t = 1$) or not ($t = 0$). Finally, D observes C's choice, and can accept the transgression, go to war over the transgression, or conduct some limited response and hassle.

The flexible-response model in Figure 1 demonstrates a setting where, when D has a better private war payoff, D will go to war less. The equilibrium likelihood of war is decreasing in type with $\underline{\theta}$ D's going to war and $\bar{\theta}$ D's hassling. Why? Within this case, mirroring the cyber-capabilities example above, we assumed that moving from $\underline{\theta}$ to $\bar{\theta}$ not only makes war a better option for D, but also makes hassling a better option for D. More specifically, while war outcomes do improve in θ, hassling outcomes improve at a faster rate in θ. This allows the shift from $\underline{\theta}$ to $\bar{\theta}$ to have the hassling payoffs surpass the wartime payoffs, thereby reducing type $\bar{D}$'s willingness to go to war. It is worth highlighting that if D lacked the hassling option, the equilibrium would follow the standard monotonicity results in Banks (1990) and Fey and Ramsay (2011), as then both types would go to war.

The model in Figure 2 demonstrates a setting where D's equilibrium payoff is lower when its private war payoff is greater. Why? In this case, as in the rally-round-the-flag example

[3] A transgression is isomorphic to an endogenous power shift.

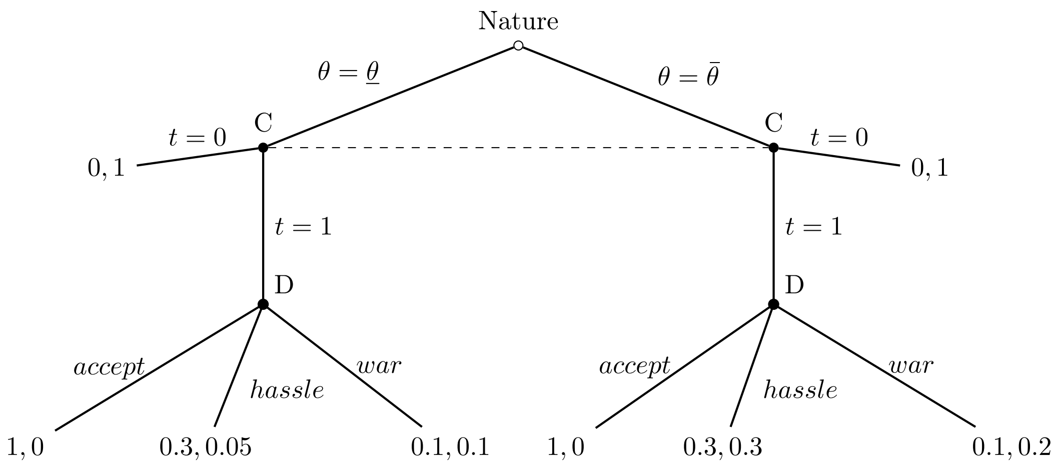

Figure 1: Greater private type θ implies less war.

C's payoffs are listed first. Here we assume that $\bar{\theta}$ has both greater wartime payoffs and hassling payoffs than $\underline{\theta}$. In equilibrium, C will transgress ($t = 1$), $\underline{\theta}$ D's will go to war and $\bar{\theta}$ D's will hassle.

above, moving from $\underline{\theta}$ to $\bar{\theta}$ not only makes war a better option for D, but also makes hassling a *worse* option for D. As a result, D's utility is decreasing in type as D's improvements in wartime payoffs fail to offset the losses from D's hassling payoffs. Again, notice that if the hassling option were unavailable, then the standard monotonicity result would hold: the equilibrium would have both types going to war, with $\bar{\theta}$ receiving a greater payoff than $\underline{\theta}$.

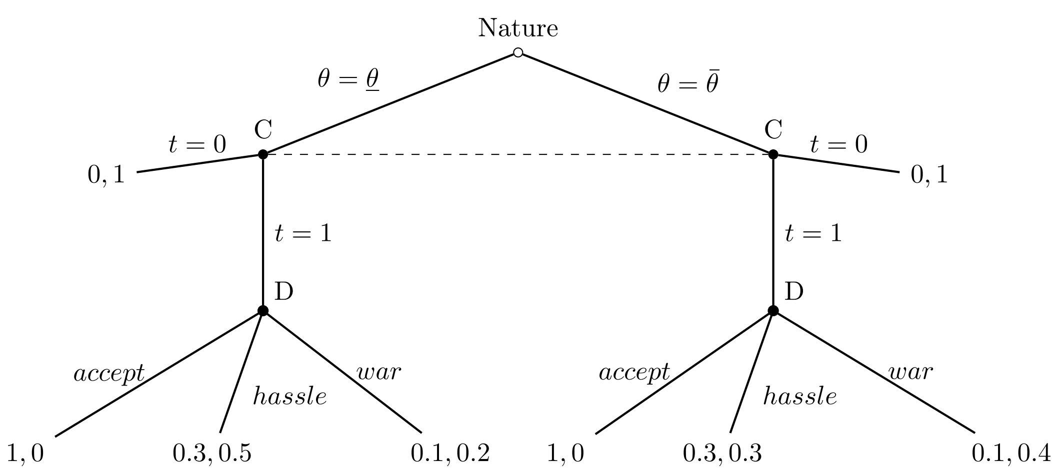

Figure 2: Greater private type θ implies lower utility

C's payoffs are listed first. Here we assume that $\bar{\theta}$ has greater wartime payoffs and lower hassling payoffs than $\underline{\theta}$. In equilibrium, C will transgress ($t = 1$), $\underline{\theta}$ D's will hassle and $\bar{\theta}$ D's will go to war.

While simple, these two models illustrate how introducing multiple related conflict options can undermine the monotonicity observed in standard crisis bargaining models. Of course, these models leave much to be desired. While it would be straightforward to create more generalized models than what is above, such an exercise would be inherently limited. A defining feature of the anarchic international order is that there is no clearly defined institution within which states interact and bargain (Waltz, 1979; Axelrod and Keohane, 1985; Niou and Ordeshook, 1990; Wendt, 1992; Mearsheimer *et al.*, 2001)—in other words, there is no clear game form. Any analysis of a game or a cluster of games has the possibility that the results are driven by the game form and would not exist if minor structural changes were made. Thus, to explain the relationship between private information and conflict in an environment with multiple related conflict options—the central aim of this paper—the game form should ideally play as little role as possible. Our mechanism design approach insures that our results are robust to the broadest set of possible institutional protocols that fall within the flexible-response crisis bargaining framework. This paper is valuable because it can save future scholars the effort of establishing (and re-establishing) the comparative static results we have here on the relationship between private type and outcomes within specific models. Instead, future work can focus on alternate facets of flexible-response crisis bargaining models, or on settings where some structure can be imposed.

Our primary contribution is a general characterization of the relationship between private war payoffs and equilibrium outcomes in flexible-response crisis bargaining games. We establish, based on model primitives and (at times) a partial equilibrium analysis, when there is a positive or negative relationship between a greater private willingness to go to war and a greater likelihood of war. At best, this means that if we know the basic properties of the defender's war payoff function and hassling cost function, we can identify the direction of the relationship between private type and war likelihood for any flexible-response crisis bargaining game. For a broad set of other cases, we can still identify the direction of the relationship, but we may also need to partially characterize equilibrium behavior. Similarly, we broadly identify when there is a positive relationship or a non-monotonic relationship between a greater private willingness to go to war and the defender's utility.

This paper makes other contributions as well. As a substantive contribution, we can offer new insights into the value of specific military technological advancements. There has now been extensive research on the value of certain coercive capabilities for specific military contexts—for example, the value of precision strike capabilities in low-level conflicts and war (Pape, 1996; Horowitz and Reiter, 2001; Kreps and Fuhrmann, 2011; Allen and Martinez Machain, 2019). Most, but not all,[4] of this research considers how these technologies fare within an active conflict or after a challenger has already transgressed. However, it is also valuable to know how these capabilities shape bargaining outcomes or the occurrence of war. Our theory can do this. Once we identify if improved private wartime capabilities improve or are detrimental to hassling capabilities—if hassling costs are increasing or decreasing in θ—then we are able to make strong statements about how changes in wartime capabilities increase or decrease utilities or the likelihood of war. Armed with our theory, future scholars can leverage past empirical and public policy research to identify how improving specific capabilities (or willingness mechanisms) can affect the outcomes in any setting where a flexible-response crisis bargaining model can apply. Additionally, as further formal advancements, we identify when "always peaceful" equilibria can arise in flexible-response crisis bargaining models, how hassling capabilities specifically shape final settlement utilities, and when to expect more (or less) hassling.

Substantively, this paper is most similar to Schram (2021), which considers multiple conflict options and a publicly observed type that systematically determines conflict utilities. Importantly, while the model in Schram (2021) does include a private type, the private type in Schram (2021) only affects hassling costs rather than affecting both hassling and war costs as is the case here. As a result, all findings here for flexible-response crisis bargaining games that break the Banks (1990) and Fey and Ramsay (2011) monotonicity results are entirely novel.[5] This paper is also similar to a range of research within the crisis bargaining framework or within the deterrence literature that considers actors with multiple possible actions within a crisis (Schultz, 2010; Powell, 2015; McCormack and Pascoe, 2017; Coe, 2018; Spaniel and Malone, 2019; **?**; Baliga *et al.*, 2020); notably, this is the first paper to systematically examine how the spillover effects of improvements in one kind of conflict capability can affect other conflict capabilities as well. This allows us to offer novel insights into a wide-range of previously unconsidered substantive settings. Methodologically, our paper is similar to a class of work on political science topics that embraces the tools of mechanism design to establish relationships between private information and outcomes. Outside of sources listed above, mechanism design has been featured in work on crisis bargaining and arbitration (Fey and Ramsay, 2009; Hörner *et al.*, 2015; Fey and Kenkel, 2019; Liu, 2021), bureaucracies and delegation (Ashworth and Sasso, 2019), firm regulation (Baron and Besanko, 1987, 1992), legislation and policy-making (Meirowitz *et al.*, 2006), voting (Aghion and Jackson, 2016),

[4]See Post (2019) as an exception.

[5]While Schram (2021) also conducts a game-form free analysis of private type, this analysis finds that a greater private willingness to hassle (or lower private costs to hassling) always results in weakly greater utilities. This is different from the U-shaped relationship between private type and utilities that can arise here (see Lemma 9).

and many others.

1 What do Flexible-Response Crisis Bargaining Models Describe?

In 2006, Israel discovered that Syria was building a nuclear reactor. Internally, Israeli decision-makers viewed the possibility of a nuclear-armed Syria as an "existential threat" to the Israeli state (Opall-Rome, 2018). In the context of the Banks (1990) and Fey and Ramsay (2011) monotonicity results—where a high private willingness to go to war over an issue leads to a greater likelihood of war—Israel's predicament was the exact setting where we might expect war to arise. But, instead of going to war, Israel used an electronic warfare attack to disable Syrian air defenses and conducted an airstrike on the reactor (Katz, 2010). Israel's airstrike against the reactor, which is known as Operation Outside the Box, successfully destroyed a critical component of the Syrian nuclear program and avoided the need for a more expansive operation.

Flexible response crisis bargaining models can capture political interactions like those surrounding the Syrian reactor development. These models formalize the following interaction. First, a challenger—State C—undertakes some opportunistic and costly action that we will refer to as a "transgression." Transgressions are politically beneficial to State C but are detrimental to State D. For example, when Syria was building a nuclear reactor in 2008, this could have eventually led to Syria possessing a nuclear bomb, which could have strengthen Syria's bargaining leverage in the future. Transgressions like this have been proposed in scholarship considering enforcement problems in bargaining (Schultz, 2010), deterrence (Fearon, 1997; Gurantz and Hirsch, 2017), or endogenous power shifts (Debs and Monteiro, 2014).[6] As examples, transgressions could be investing in conventional, nuclear, space, or cyber military technologies (Debs and Monteiro, 2014; Gartzke and Lindsay, 2017; Spaniel, 2019), forming alliances (Benson and Smith, 2020), or securing geopolitically valuable territory (Fearon, 1996; Powell, 2006).

In response to the transgression, a defender—State D—has the choice between accepting the transgression through a peaceful settlement, going to war to decisively resolve the political issues between the states, or in engaging in some low-level actions that undercut the

[6] The "transgression" here is similar to how a large literature treats the challenger's first move in a deterrence game. See Huth (1999) for a review of early work on deterrence, as well as Kydd and McManus (2017); Smith (1998); Di Lonardo and Tyson (2017); Chassang and Miquel (2010); Baliga and Sjöström (2020).

transgression. The first two options (a peaceful settlement or war) are standard in the crisis bargaining or deterrence frameworks. We will refer to the the last option (D's low-level response) as "hassling." As originally defined in **?**, hassling is the use of limited conflict to degrade a challenger's rise. Our use of the term here is consistent with this definition, but also refers to any actions by the defender that undercuts the challenger's transgression. In the Syria case, Israel detected Syria's nuclear reactor and destroyed it. This would be consistent with hassling because it was a destructive blow to Syria's nuclear program, but it was not a decisive military move that would prevent the Assad regime from every possessing a nuclear weapon.[7] Hassling can take the form of limited airstrikes (Reiter, 2005; Fuhrmann and Kreps, 2010), hybrid conflict (Lanoszka, 2016), aspects of gray-zone conflict (Mazarr, 2015; Gannon *et al.*, 2020), (limited) preventive war (Levy, 2011), sanctions (McCormack and Pascoe, 2017), or arming (Coe, 2018).

Within these interactions, we assume that D possesses a private type that influences both war and hassling capabilities. In other words, if state D is (privately) very capable at conducting war or very willing to conduct war, we might also expect that state D is systematically better (or worse) at conducting hassling. To offer perhaps the simplest example, suppose a state is privately very hawkish on the matter of its neighbors developing nuclear weapons. Based on these private preferences, this state might be very willing to conduct a war and more willing to implement a low-level hassling attack to weaken a target. Ultimately, whether better private capabilities or a greater willingness to go to war also makes low-level options more or less appealing is ultimately an empirical question; we discuss this later in section 5.

2 The Flexible Response Model

Here we present the flexible-response crisis bargaining framework. In this framework, negotiations may end efficiently in peace or inefficiently either in war or in one of a continuum of possible inefficient non-war outcomes. We assume a state's private information may affect its payoffs from both kinds of inefficient outcomes, and we use the tools of Bayesian mechanism design to obtain general results about the relationship between private types and the equilibrium properties of flexible-response crisis bargaining games.

[7] To offer an example of a decisive military move, the 2003 U.S. invasion of Iraq decisively prevented the Ba'athist regime from attaining nuclear weapons by overthrowing the Ba'athist regime. This invasion is consistent with this model's treatment of war.

2.1 Structure of the Interaction

At the outset of the interaction, Nature assigns D's type, $\theta \in \mathbb{R}$. Without loss of generality, we assume D's war payoff increases with D's type,[8] while D's cost of hassling may increase, decrease, or neither. The realized value of θ is known only to D, but the prior distribution from which it is drawn is common knowledge. Let F denote the CDF of this prior distribution, and let Θ denote its support.

The interaction between the states takes the familiar form of a crisis bargaining game, except each state may engage in activity that affects outcomes short of war. First, C selects transgression $t \in \mathcal{T} \subseteq \mathbb{R}_+$. Following this choice, C and D partake in a bargaining process that may end in war or in some hassling response by D. Like Banks (1990) and Fey and Ramsay (2011), we place no particular structure on the bargaining process. We simply assume that each player chooses from a set of available bargaining actions; these choices determine whether the game ends in war, and, if not, how the prize is divided. In the negotiation stage, we let $b_C \in \mathcal{B}_C$ denote C's bargaining strategy (offers, counteroffers, accept-reject plans, etc.). D's strategy consists of an analogous bargaining strategy $b_D \in \mathcal{B}_D$, as well as a level of hassling, $h \in \mathcal{H} \subseteq \mathbb{R}_+$.[9] A *game form* G consists of the bargaining action spaces, $\mathcal{B}_A$ and $\mathcal{B}_D$, along with an *outcome function* g mapping from the choices (t, h, b_C, b_D) into the set of possible crisis bargaining outcomes.[10] We decompose the outcome function g into three components: whether war occurs, what C receives from bargaining, and what D receives from bargaining.[11] Whether war occurs or not depends solely on actions taken in bargaining. Let $\pi^g(b_C, b_D) \in \{0, 1\}$ be an indicator for whether the interaction ends without war.[12] Conditional on war not occurring, each player's payoff depends on the bargaining behavior, C's choice of transgression, and D's selection of hassling. Let $V_C^g(t, h, b_C, b_D)$ and $V_D^g(t, h, b_C, b_D)$ denote the benefits that C and D receive, respectively, in case war is avoided.

Payoffs depend on the outcome of bargaining, including the costs of the hassling or transgres-

[8] In terms of the traditional "costly lottery" formulation of crisis bargaining games, greater θ may correspond to a greater probability of winning, a lower cost of fighting, or both. However, we do not impose a costly lottery model—all that matters for our purposes is that D's expected utility from war strictly increases with θ.

[9] We place no restriction on whether hassling is chosen before, during, or after the bargaining process—all that matters for our purposes is that the cost of any given level $h \in \mathcal{H}$ is independent of b_C and b_D.

[10] The game form represents the elements of the model that are specific to a particular bargaining protocol. Implicitly, then, we take the type space, prior distribution, transgression action set, hassling action set, cost functions, and war payoff functions as primitives of the model rather than features of a game form G.

[11] Unlike in Banks (1990) and Fey and Ramsay (2011), we allow for inefficient settlements. This means D's value from bargaining cannot be immediately deduced from C's, and vice versa.

[12] By ruling out $\pi^g \in (0, 1)$, we are implicitly assuming the bargaining process has no exogenous random components (see Fey and Kenkel, 2019).

sion when war does not occur.[13] War payoffs may depend on D's private information, but do not depend on any of the endogenous choices in the game, including transgressions and hassling.[14] We therefore write war payoffs as $W_A(\theta)$ and $W_D(\theta)$. We assume W_D is strictly increasing, so higher types of D can be interpreted as stronger in wartime. If war is avoided, each player receives their division of the spoils but must pay the cost of their transgression or hassling. Let $K_C(t, h)$ denote the cost to A, and let $K_D(h, \theta)$ denote the cost to D. We assume that K_C is strictly increasing in t and weakly decreasing in h, and we assume K_D is strictly increasing in h. We let $h = 0$ denote no hassling, which entails assuming $0 \in \mathcal{H}$ and $K_D(0, \theta) = 0$ for all θ. Putting these together, the players' utility functions in a given game form are as follows:

$$u^g_D(t, h, b_C, b_D \,|\, \theta) = (1 - \pi^g(b_C, b_D))W_D(\theta) + \pi^g(b_C, b_D)[V^g_C(t, h, b_C, b_D) - K_C(t, h)], \quad (1)$$
$$u^g_D(t, h, b_C, b_D \,|\, \theta) = (1 - \pi^g(b_C, b_D))W_D(\theta) + \pi^g(b_C, b_D)[V^g_D(t, h, b_C, b_D) - K_D(h, \theta)]. \quad (2)$$

We restrict our attention to game forms in which neither player can force a settlement on the other. This assumption reflects the anarchic nature of international politics, in which states always have the option to resort to force if desired. A sufficient condition is that each player has an action $b_i \in \mathcal{B}_i$ such that $\pi^g(b_i, b_j) = 0$ (war is guaranteed) for all $b_j \in \mathcal{B}_j$. As we show below, this condition places important limits on what kinds of outcomes are sustainable as equilibria.

The relationship between D's private type and hassling ability is central to our analysis. We will show that the effects of θ on equilibrium outcomes depend critically on whether greater payoffs from war are associated with higher or lower costs of hassling. In general, when comparing types θ' and θ'', we say that θ'' has greater hassling effectiveness than θ' if $K_D(h, \theta'') < K_D(h, \theta')$ for all $h > 0$. We focus on cases where there is a monotone relationship between private type and hassling effectiveness. We say θ *improves hassling effectiveness* if $K_D(h, \theta'') < K_D(h, \theta')$ for all $h > 0$ and $\theta' < \theta''$. In the opposite case, when K_D strictly decreases with D's type, we say θ *degrades hassling effectiveness*.

Our definition of hassling effectiveness concerns the relationship between D's type and the absolute cost of hassling. We can obtain stronger results if we go further, relating D's type to the *marginal* cost of hassling. We say the hassling cost function satisfies *decreasing differences* if types with lower absolute costs also have lower marginal costs. This condition

[13] We restrict attention to models in which the players' payoffs from any peaceful outcome do not depend on D's private information, except insofar as that private information affects the cost to D of the chosen hassling level h.

[14] We relax this in an extension below, in which transgressions and hassling may affect war payoffs.

ensures that D's settlement utility has the single-crossing property, allowing us to characterize monotone comparative statics without imposing strong functional forms (Ashworth and Bueno de Mesquita, 2006).

Definition 1. The cost function K_D has *decreasing differences* in h and θ if

$$\begin{aligned}&\theta' \text{ has greater hassling effectiveness than } \theta \Rightarrow \\ &\quad K_D(h',\theta') - K_D(h,\theta') < K_D(h',\theta) - K_D(h,\theta) \text{ for all } h < h'.\end{aligned} \tag{DD}$$

2.2 Solution Concept and Direct Mechanisms

We restrict attention to pure strategy perfect Bayesian equilibria of each flexible-response crisis bargaining game. Depending on the bargaining protocol and the equilibrium selected, the equilibrium path may be very complex, involving numerous offers and counteroffers before concluding, or it may be simple, ending quickly in war or a settlement. We will not dwell on the details of bargaining itself, as our primary concern is the *outcome* of the interaction: whether war prevails, and if not, what each party receives from a bargained outcome. Each player's bargaining behavior only affects their payoffs insofar as it affects these components of the outcome.

We will focus on the incentives of D, the player with private information. Given an equilibrium of a flexible-response crisis bargaining game, we can summarize the outcome of the game for each type of D with three functions:[15]

- Their hassling level, $h(\theta)$.
- Whether a bargained outcome prevails, $\pi(\theta)$.
- Their settlement value in case of a bargained resolution, $V_D(\theta)$.

A *direct mechanism* for D consists of these functions, (h, π, V_D). If type θ of D were to follow the equilibrium bargaining strategy of type θ', D's expected utility from doing so would be:

$$\Phi_D(\theta' \,|\, \theta) = \underbrace{(1-\pi(\theta'))W_D(\theta)}_{\text{war}} + \underbrace{\pi(\theta')[V_D(\theta') - K_D(h(\theta'),\theta)]}_{\text{bargained outcome}}.$$

While mimicking another type's strategy may change the hassling level, the occurrence of conflict, and the settlement value for θ, it does not change D's war payoff, nor the cost D

[15] In the Appendix, we formally define an equilibrium and describe how a direct mechanism can be derived from it. See the discussion in Fey and Ramsay (2011).

pays for any given hassling level.[16] In equilibrium, no type may have a strict incentive to mimic another type's bargaining strategy. We can phrase this requirement as an *incentive compatibility* condition on the direct mechanism. Let $U_D(\theta)$ denote each type's expected utility along the path of play, so that $U_D(\theta) = \Phi_D(\theta \,|\, \theta)$.

Definition 2. A direct mechanism (h, π, V_D) is *incentive compatible* if

$$U_D(\theta) \geq \Phi_D(\theta' \,|\, \theta) \qquad \text{for all } \theta, \theta' \in \Theta. \tag{IC}$$

To identify regularities in the outcomes of flexible-response crisis bargaining games, we will analyze incentive compatible direct mechanisms. We rely on the *revelation principle*: for any Bayesian Nash equilibrium of a particular game form, there is an incentive compatible direct mechanism that yields the same outcome (Myerson, 1979). Logically, this means that if we find that some property holds for all incentive compatible direct mechanisms, than it is true of all equilibria of all flexible-response crisis bargaining games. Without bogging ourselves down in the particulars of how crisis bargaining plays out in any particular game, we are still able to characterize robust properties of the outcomes of any flexible-response crisis bargaining game.

Recall that we only consider game forms in which neither player can impose a settlement on the other. This condition ensures that no type of D may receive less than its war payoff in equilibrium—if a settlement would yield less, then it would be profitable for D to deviate to fighting. This requirement amounts to a participation constraint in the language of mechanism design, or what Fey and Ramsay (2011) call voluntary agreements in the crisis bargaining context.

Definition 3. A direct mechanism (h, π, V_D) has *voluntary agreements* if

$$\pi(\theta)[V_D(\theta) - K_D(h(\theta), \theta)] \geq \pi(\theta) W_D(\theta) \qquad \text{for all } \theta \in \Theta. \tag{VA}$$

Naturally, the voluntary agreements condition is automatically satisfied for those types that fight on the path of play. The constraint only applies to the types that settle—the settlement

[16]The definition of Φ_D illustrates an important difference between the flexible-response framework and the environment studied by Fey and Ramsay (2009), who also allow for inefficient bargained settlements. In their model, the inefficiency loss due to reporting θ' is the same for all types θ. By contrast, we assume the cost of hassling is a function of the true type θ, which implies that different types may value the "same" settlement differently.

must yield at least as much as their war payoff, even when accounting for the costs of the hassling. Throughout the analysis, we will restrict attention to direct mechanisms that satisfy both (IC) and (VA), as any equilibrium of a flexible-response crisis bargaining game with voluntary agreements must be outcome-equivalent to some such mechanism (Fey and Ramsay, 2011).

3 Private Type and the Probability of War

In crisis bargaining games without flexible responses, private signals of high strength or resolve are associated with a greater equilibrium probability of war. A simple intuition drives this result: if some type finds it worthwhile to run a high risk of war to receive a better deal at the bargaining table, then all stronger types must be willing to run at least as great a risk. For stronger types, the benefits from any settlement are the same, while their payoffs from the fail case of fighting are better.

For equilibria with no hassling, our model recovers the classic monotone relationship between private type and the likelihood of conflict. As long as $h(\theta) = 0$ for all types, stronger types never have a lower probability of conflict than weaker types. The following result recovers Lemma 1 of Banks (1990) as a special case in our environment.[17]

Lemma 1. *If* $h = 0$ *and* $\theta' < \theta''$*, then* $\pi(\theta') \geq \pi(\theta'')$*.*

This result confirms that the exceptions we find to the classic model are due exclusively to our introduction of flexible responses, not to any other feature of our modeling environment. That said, once states begin to employ alternative instruments for altering the balance of bargaining power, however, this straightforward relationship between private strength and the risk of war holds only under special conditions.

The flow chart in Figure 3 summarizes our findings on the relationship between private information and the occurrence of conflict. If private strength degrades hassling effectiveness, then stronger types are more likely to fight, just as in traditional crisis bargaining models. The same is true if private strength increases hassling effectiveness and has a stronger effect on war payoffs than on the cost of hassling, or, formally, if WURI holds. We find that weaker types are more likely to fight—the opposite of the traditional result—when private strength improves hassling effectiveness but affects war payoffs less than hassling costs, or

[17] Though we restrict to deterministic outcomes and thus $\pi \in \{0, 1\}$ in the bulk of our analysis, the proof of Lemma 1 holds even if we allow for probabilistic outcomes. All proofs are in the Appendix.

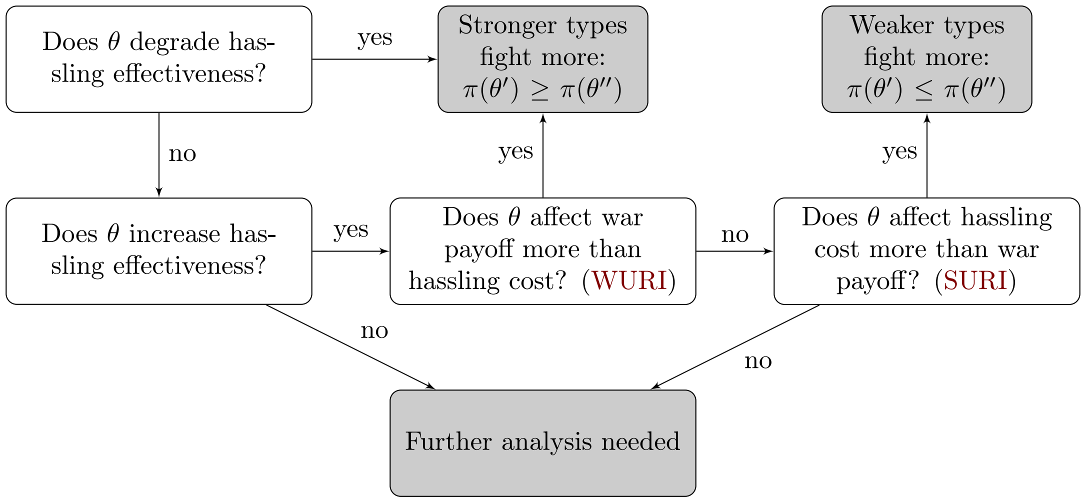

Figure 3: The relationship between type ($\theta' < \theta''$) and war likelihood in all flexible-response crisis bargaining games.

when, formally, SURI holds If there is no monotone relationship between private strength and the cost of hassling or if, when stronger types are more effective, neither WURI nor SURI hold, then we cannot pin down a clear relationship with the probability of conflict.

When hassling takes place, the effect of private type on the chance of war depends on whether a state's private war capability improves or degrades its hassling effectiveness. If private type degrades hassling effectiveness—i.e., if types with greater war payoffs also have greater costs of hassling—then higher types are weakly more likely to go to conflict.

Proposition 1. *Assume θ degrades hassling effectiveness. If $\theta' < \theta''$, then $\pi(\theta') \geq \pi(\theta'')$.*

Because the value of an efficient settlement without hassling is the same regardless of θ, higher types have the most incentive to choose war over an efficient settlement. If private type degrades hassling effectiveness, then this logic carries over to settlements involving hassling as well. If some type of D prefers war over a settlement with hassling level $h \geq 0$, then all stronger types must have the same preference: they have an even higher war payoff and would pay a greater cost from the same level of hassling.

When private type is instead associated with lower costs of hassling, we need even more conditions to characterize its relationship with the equilibrium occurrence of war. In this case, not only is war more attractive to stronger types of D, but so is any given settlement with hassling. Because war payoffs and settlement payoffs are now moving in the same direction as a function of D's type, the critical question for our purposes is which rate of increase is quicker. The probability of conflict increases with D's private strength in equilibrium if θ has a stronger effect on war payoffs than on the cost of hassling. On the other hand, if the effect of θ on hassling cost is dominant, then weaker types are more likely to fight in equilibrium—the opposite of the pattern in traditional crisis bargaining games.

To prove these claims, we must formally state what it means for θ to affect war payoffs more than the cost of hassling, or vice versa. We say the war utility is relatively increasing (WURI) when θ has a greater marginal effect on war payoffs than on the cost of hassling. In the opposite case, we say the settlement utility is relatively increasing (SURI).

Definition 4. In a direct mechanism, the *war utility is relatively increasing* if

$$\begin{gathered} W_D(\theta'') - W_D(\theta') > K_D(h(\theta''), \theta') - K_D(h(\theta''), \theta'') \\ \text{for all } \theta', \theta'' \in \theta' \text{ such that } \theta' < \theta'' \text{ and } \pi(\theta'') = 1. \end{gathered} \tag{WURI}$$

The *settlement utility is relatively increasing* if

$$\begin{gathered} W_D(\theta'') - W_D(\theta') < K_D(h(\theta'), \theta') - K_D(h(\theta'), \theta'') \\ \text{for all } \theta', \theta'' \in \theta' \text{ such that } \theta' < \theta'' \text{ and } \pi(\theta') = 1. \end{gathered} \tag{SURI}$$

If either of these holds and private strength increases hassling effectiveness, then we can identify a monotonic relationship between private information and the equilibrium chance of conflict for any flexible-response crisis bargaining model.

Proposition 2. *Assume θ improves hassling effectiveness, and let $\theta' < \theta''$. If* (WURI) *holds, then $\pi(\theta') \geq \pi(\theta'')$. If* (SURI) *holds, then $\pi(\theta') \leq \pi(\theta'')$.*

This result shows that the conventional relationship between private information and the likelihood of conflict is not robust to the introduction of hassling that affects payoffs from bargaining. Assuming that types with greater battlefield effectiveness are also more effective at hassling activities, the relationship between θ and the likelihood of conflict depends critically on the technology of hassling. If the marginal effect of D's type on the costs of hassling always outweighs its effect on the war payoffs, then we have the opposite of the usual result, with stronger types less likely to fight on the path of play.

The two conditions we have outlined here are mutually exclusive (except in the trivial case where all types end up fighting in equilibrium), but they are not mutually exhaustive. Depending on the functional forms of W_D and K_D, it is possible for the marginal effect of θ on the war payoff to be relatively strong for some types and relatively weak for others. If so, we cannot generally characterize the relationship between D's private type and which outcome prevails in equilibrium.

While Proposition 2 is useful for understanding how private information affects the occurrence of war in flexible response crisis bargaining games, its practical applicability is somewhat limited. Ideally, we would be able to say on the basis of the model primitives—the war payoff and hassling cost functions—whether stronger types will be associated with a greater likelihood of conflict in any given strategic environment. However, the WURI and SURI conditions do not exclusively concern model primitives, as they depend on the levels of hassling chosen on the path of play. This raises the possibility that the relationship between D's private type and the likelihood of conflict may vary depending on the exact bargaining protocol.

With additional conditions on the model primitives, we can ensure that the war utility is relatively increasing, meaning the likelihood of conflict increases with D's type. In particular, we need the cost function to have decreasing differences and for the marginal effect of θ on the war payoff to always exceed its marginal effect on the hassling cost when h is at its upper bound. Under these conditions, higher types are more likely to end up at conflict in the equilibria of all flexible response crisis bargaining games, regardless of the exact negotiating protocol employed.

Lemma 2. *Assume θ improves hassling effectiveness, (DD) holds, and $\max \mathcal{H} = \bar{h} < \infty$. If $W_D(\theta'') - W_D(\theta') > K_D(\bar{h}, \theta') - K_D(\bar{h}, \theta'')$ for all $\theta', \theta'' \in \theta'$ such that $\theta' < \theta''$, then (WURI) holds.*

There is not an analogous sufficient condition for the settlement utility to be relatively increasing. The obstacle to such a condition is our assumption that $h = 0$ is always feasible at zero cost.[18] This means the marginal effect of θ on the cost of the hassling is zero for $h = 0$, ruling out any kind of sufficient condition for the marginal effect of θ on the hassling cost to always exceed its effect on the war payoff. At most, if we assume decreasing differences in the cost of hassling, we can make the SURI condition slightly less onerous to check. In this case, letting $\underline{h}$ denote the minimal hassling among types that end up in a bargained resolution, a sufficient condition is that $W_D(\theta') - W_D(\theta) < K_D(\underline{h}, \theta) - K_D(\underline{h}, \theta')$ for all $\theta < \theta'$. If this

[18] The normalization of the cost to zero is immaterial, but the constancy of the cost of across types of D is important here.

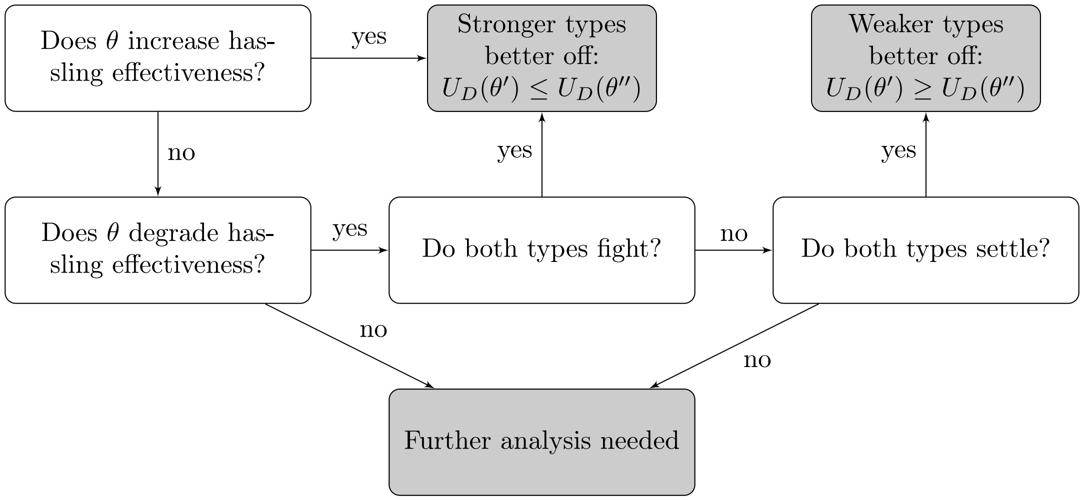

Figure 4: The relationship between type ($\theta' < \theta''$) and equilibrium payoffs in all flexible-response crisis bargaining games.

condition holds, then the equilibrium must entail low types of D fighting and high types of D settling.

4 Private Type and Payoffs

A second key regularity in crisis bargaining games without flexible responses is that a better private war capability corresponds to a better equilibrium payoff, even when war does not occur. Greater private strength increases an actor's willingness to run the risk of war. Risky bargaining actions thus serve as signals of strength, in turn allowing the actor to extract more from negotiations in case the interaction does not ultimately end in war.

We recover the same relationship between private strength and equilibrium payoffs in the flexible response environment when hassling does not take place along the path of play. The following result is our analogue of Lemma 4 from Banks (1990).

Lemma 3. *If $h = 0$ and $\theta' < \theta''$, then $U_D(\theta') \leq U_D(\theta'')$.*

Once we introduce additional ways to alter the balance of spoils, the tight relationship between private war capability and equilibrium payoffs no longer holds. The flowchart in Figure 4 summarizes the relationship between private strength and utility in flexible response crisis bargaining games. Obviously, among types that end up fighting in equilibrium, stronger ones are always better off. Besides that, however, private strength is only guaranteed to increase payoffs when it is also associated with greater hassling capability. If private strength instead degrades hassling capability, we find the opposite of the usual result—types with lower private strength have greater payoffs in case of peace.

We recover the positive relationship between private strength and equilibrium payoffs when θ improves hassling effectiveness. In this case, stronger types have an advantage in both channels of bargaining leverage—hassling and the threat of war—and therefore never come away worse off at the bargaining table. In fact, a stronger type has a strictly greater payoff than all weaker types whenever it goes to war in equilibrium or it engages in non-zero hassling.

Proposition 3. *Assume θ improves hassling effectiveness. If $\theta' < \theta''$, then $U_D(\theta') \leq U_D(\theta'')$. The inequality is strict if $\pi(\theta') = 0$ or $h(\theta') > 0$.*

A simple incentive compatibility logic is behind this result. Consider the outcomes for two types of D, one weaker and one stronger. If the weaker one goes to war, then obviously the stronger one could do better by going to war as well. Conversely, if the weaker one settles, the stronger type could get the same terms of settlement as the weaker one by choosing the same hassling level and bargaining actions. Moreover, under the condition that private type improves hassling capabilities, the stronger type would pay no greater a cost (in fact, strictly less if $h > 0$) for the hassling. No matter what the outcome for the weaker type, the stronger type has a bargaining strategy available that results in a weakly better payoff. Consequently, the stronger type's equilibrium payoff must be no lower than the weaker one's.

If private strength is instead associated with lower hassling effectiveness, then we find an exception to the traditional positive relationship between private type and equilibrium payoff. In this case, the relationship is U-shaped. Low types, which have poor battlefield effectiveness but relatively low costs of hassling, choose to settle rather than to fight in equilibrium. Among these types, lower private strength is associated with greater hassling ability, and thus a greater equilibrium payoff. At a certain level of strength, however, it becomes profitable to fight rather than settle. After this point, greater military strength by necessity leads to a greater payoff.

Proposition 4. *Assume θ degrades hassling effectiveness. There exists $\hat{\theta}$ such that $\pi(\theta) = 1$ for all $\theta < \hat{\theta}$ and $\pi(\theta) = 0$ for all $\theta > \hat{\theta}$. If $\theta' < \theta'' < \hat{\theta}$, then $U_D(\theta') \geq U_D(\theta'')$ (strictly if $h(\theta'') > 0$). If $\hat{\theta} < \theta' < \theta''$, then $U_D(\theta') < U_D(\theta'')$.*

This result illustrates the new sources of bargaining leverage that arise in crisis bargaining games with flexible responses. In ordinary crisis bargaining games, a state's sole source of bargaining power is its threat to resort to war. In our framework, hassling provides another means of shifting the peacetime balance of spoils. There is a direct effect of hassling efficiency, where more effective types end up better off because they can afford to hassle more, thereby shifting the balance of goods in their favor. There is also an indirect effect: more effective types can pay a lower cost to undertake the same amount of hassling. Both effects contribute to the negative relationship between private strength and payoffs from peace when θ degrades hassling effectiveness.

In analyzing private strength's relationship with both the probability of war and equilibrium payoffs in flexible response crisis bargaining games, we have identified conditions where our results diverge with the traditional patterns and other conditions where they agree. Interestingly, we find only one case in which *both* the chance of conflict and the equilibrium utility increase with private strength as usual—namely, when θ increases hassling effectiveness but has an even stronger effect on war payoffs (WURI). The ordinary crisis bargaining model may, in a sense, be considered a special case of these more general conditions. If private strength increases hassling effectiveness while having a relatively low effect on war payoffs (SURI), then equilibrium payoffs increase with θ as in traditional models, but the probability of war decreases. We see the converse pattern if private strength degrades hassling effectiveness: the probability of war increases with θ as usual, but the equilibrium utility is U-shaped.

5 Empirical Implications

What are the empirical implications of our theory? Consider the ongoing debate over the usefulness of aerial bombing in conflict. To quickly summarize a subset of findings, existing research has shown that aerial bombings or precision strike capabilities can be useful in hassling operations (Kreps and Fuhrmann, 2011), can be useful in a conventional war (Pape, 1996; Horowitz and Reiter, 2001; Allen and Martinez Machain, 2019), and may not be as useful in a counterinsurgency (Lyall, 2013; Dell and Querubin, 2018). While this empirical research agenda is quite deep, the existing literature does not speak to, for example, what precision strike technology means for shaping deterrence or bargaining.[19] In every paper cited above, the value of these weapons were evaluated conditional upon a bargaining failure having occurred. It would also be valuable to know how these weapons influence decision

[19] One notable exception is Post (2019), which analyzes airpower events (i.e. military exercises, mobilizations, shows of force, deployments of military assets, or other military signal) as signals for use in compellence.

making in the lead-up to crises, whether they lead to more (or less) war, and whether they produce expected windfalls for the state that developing the bombing capabilities. But, these latter questions are difficult to answer empirically; identifying or constructing a counterfactual world where a state did not develop precision strike capabilities is challenging, as is identifying how these capabilities (or their absences) shape international crisis initiation or behavior. Our theory is well suited to address these questions. The flowcharts in Figures **??** and **??** demonstrate that once we identify if improved private wartime capabilities improve or are detrimental to hassling capabilities—if hassling costs are increasing or decreasing in θ—then we are able to make strong statements about how changes in wartime capabilities increase or decrease utilities or the likelihood of war.

To illustrate how this could work, considered a crisis where (a) a defender state possesses private information about their ability to conduct precision strikes, (b) the most feasible low-level conflict option is a precision strike (along the lines of Operation Desert Fox), and (c) the war option would be a conventional war (along the lines of Operation Desert Storm). Because precision strike technology is dual-use for both hassling and war, then these private capabilities would have a positive effect on both war and hassling outcomes (hence the " + " symbols). Thus, if (a)-(c) held, as the flowcharts in figures **??** and **??** show, any flexible-response crisis bargaining model of this case would find that a defender with a better private ability to conduct precision strikes would end up with greater final payoffs, but not necessarily a greater likelihood of engaging in a conventional war. On the other hand, consider a crisis where (a) and (b) held, but, instead of (c), it was the case that (d) the war option would be a protracted counterinsurgency, where precision strike capabilities may be less effective (hence the " - " symbol).[20] If (a), (b), and (d) held, then any flexible-response crisis bargaining model of this case would find that a defender state with a better private ability (or private willingness) to conduct precision strikes would be less likely to enter into a war, and may or may not end up with greater final expected payoffs.[21]

A number of other conflict capabilities that are dual-use for hassling and war. During Operation Outside the Box (2007), Israel disabled Syrian air defenses with an electronic warfare (EW) attack. While the full details of the EW attack are not disclosed, an attack that allowed multiple Israeli aircraft to enter Syria and conduct a raid without harassment plausibly could have been used to conduct a more extensive conventional attack as well (Katz, 2010). Additionally, assorted cyberattacks have been used both independently (Stuxnet and Esto-

[20] Alternatively, in a setting with budget restrictions, it is plausible that investments in precision strike capabilities were made at the expense of counterinsurgency capabilities.

[21] While throughout the text we treat θ as improving war capabilities, we can draw the conclusions stated here due to 1.

nian cyber attacks), as a part of a cluster of operations against a target state (the NotPetya attacks targeting Ukraine), or as part of a conventional war (Russia-Georgia War, 2008) (Buchanan, 2020; Gannon *et al.*, 2020). Furthermore, developments in anti-satellite technologies have opened the possibility for disruption of GPS signals, which could both create low-level disruptions and create serious problems for modern air and sea warfare (Harrison *et al.*, 2020). While we do not have the space to discuss every dual-use technology case at length, a number of other capabilities positively affect hassling and war in straightforward ways: the continued development of airlift capabilities can facilitate special operations for use in low-level conflict, conventional war, and irregular warfare (Bolkcom, 2007; Pietrucha and Renken, 2019); coordinating with or providing support to violent non-state actors found use in hassling and war (Schultz, 2010; **?**); and an ability to control the civilian information environment have been used in hassling operations (Russia in Eastern Ukraine, 2015-on) and within hearts-and-minds counterinsurgencies.

Of course, a number of capabilities are less effective in some forms of conflict. Just as precision strike capabilities may not be well suited for a protracted hearts-and-minds COIN campaign (Lyall, 2013; Dell and Querubin, 2018), the electronic, cyber, or anti-satellite tools can be very expensive and may not be as well suited for a COIN environment;[22] to the extent that, for example, states over-invest in precision strike capabilities, those states may be less able to conduct a war requiring extensive COIN operations. Additionally, while civilian information operations can help in a hassling or COIN context, these operations may not be as well-suited for a conventional conflict. As another example, the U.S. Navy, Marine Force, and Coast Guard has jointly issued a report suggesting that conventional navy vessels designed for legality in conventional warfare may not be as effective in deft handling of gray zone attacks in the pacific theater, and a more agile fleet could better serve in "gray zone" operations (Berger *et al.*, 2020; Owen, 2021).

Outside of developing capabilities, other factors could shape a leadership's incentives in such a way that alters their willingness to engage in some forms of conflict over others. Statistical analyses of the "rally-round-the-flag" phenomenon suggest that larger military operations, especially wars, generate an increase in public support for domestic leadership (Baker and Oneal, 2001; Chapman and Reiter, 2004). This effect could alter a leader's preferences to prefer war to hassling operations. Alternatively, domestic political economy considerations can shape the incentives of leaders. In 1954 the United States provided arms, funds, and training to Guatemalan rebels who overthrew Jacobo Árbenz and installed right-wing dictator Castillo Armas; this move benefited the politically connected United

[22] There are some notable exceptions, see Shachtman (2011).

Capability/Willingness Mechanism	**Hassling Type**	**War Type**	**Rationale**
Precision Strikes/Drones	Targeted strikes against military facilities (+)	Conventional War (+)	Dual-Use Technology
Electronic/Cyber/Anti-Satellite Attacks	Disruption of military systems or weapons facilities (+)	Conventional War (+)	Dual-Use Technology
Airlift capabilities	Special operations (+)	Conventional/ Irregular War (+)	Dual-Use Technology
Militants on Retainer	Supporting low-level conflict/insurgency (+)	Conventional/ Irregular War (+)	Dual-Use Technology
Domestic Information Operations	Undermining domestic authority/promoting discord (+)	Irregular War (+)	Dual-Use Technology
Domestic Information Operations	Undermining domestic authority/promoting discord	Conventional War (-)	Specialized Technology/Budgetary Issues
Precision, Electronic, Cyber, Anti-Satellite Attacks	Targeted strikes against military facilities (+)	Irregular War/COIN (-)	Specialized Technology/Budgetary Issues
Conventional Navy gray hulls optimized for high-end naval warfighting	Countering naval gray zone operations (-)	Conventional War (+)	Specialized Technology/Budgetary Issues
Domestic Electoral Considerations	Sanctions, hassling, or any other low-level response (-)	Conventional War (+)	Rally 'round the Flag
Domestic Political Economy Considerations	Sanctions or low-level conflict (+/-)	Conventional War (+/-)	Political Elite Disconnect

Table 1: How various capabilities or willingness mechanisms affect hassling and war capacities. The "+" symbol ("-" symbol) indicates that the capability or willingness mechanism improves (is detrimental to) the listed hassling operations or war type.

Fruit Company (Kinzer, 2007, 125–147) in ways that sanctions would not. Similarly, while a "blood for oil" hypothesis may not fully explain 2003 Iraq invasion (Paul, 2003; Stokes, 2007), it could have still shaped a private willingness for U.S. leadership to go to war rather than hassle. Of course, this is not to say that relevant domestic actors always prefer war, as the use of international sanctions or tariffs to weaken a regime can also create domestic winners and losers.

6 Additional Analysis

6.1 Possibility of Peace

Previous game-free analyses of crisis bargaining have sought to identify sufficient conditions for peace to be an equilibrium outcome (Fey and Ramsay, 2009, 2011). The mechanism design approach is well suited for studying the possibility of peace. We may ask whether there is *any* crisis bargaining game that would lead to peace as an equilibrium outcome. Any given game form may result in a positive probability of war, but that does not mean war is inevitable—unlike bargaining in legislatures, states are not bound to follow any particular protocol.

When working with flexible-response crisis bargaining games, we must be explicit about what it means for an equilibrium to be peaceful. At a minimum, as in ordinary crisis bargaining games, the game must end with a negotiated settlement for all types $\theta \in \Theta$. Furthermore, because transgressions and hassling may be interpreted as forms of low-level conflict, we will focus on equilibria in which C chooses $t = 0$ and each type of D chooses $h = 0$. Mirroring the terminology of Fey and Ramsay (2011), we will call an equilibrium meeting these conditions *always peaceful*.

In our baseline flexible-response context, the sufficient condition for peace is virtually the same as in ordinary crisis bargaining models. In particular, it must be possible to divide the pie so as to simultaneously satisfy both C, assuming C's knowledge of D's type is limited to the prior distribution, and the strongest type of D. In what follows, let $\hat{W}_C = \mathbb{E}[W_C(\theta)] = \int_\Theta W_C(\theta)\, dF(\theta)$, C's prior expectation of its own war payoff.

Proposition 5. *If $\hat{W}_C + W_D(\bar{\theta}) \leq 1$, then there is a flexible-response crisis bargaining game form with voluntary agreements that has an always peaceful equilibrium.*

The condition of Proposition 5 is least likely to hold when the distribution of D's type is right-skewed. In this case, C's expected war payoff will be relatively high, since D's type is likely to be low. It will thus be impossible to satisfy the (rarely occurring) strongest type of D while giving C at least its expected war payoff.

If C's war payoff is independent of D's type (i.e., D's type only affects its cost of fighting, not its probability of victory), then the condition of Proposition 5 is sure to hold. A distribution of the pie following the probability of war will be acceptable both to C and to all types of D. The following result is a direct analogue of Proposition 2 in Fey and Ramsay (2011).

Corollary 1. *If $W_C(\theta) = p - c_C$ and $W_D(\theta) = 1 - p - c_D(\theta)$, where $c_D : \Theta \to \mathbb{R}_+$, then there is a flexible-response crisis bargaining game form with voluntary agreements that has an always peaceful equilibrium.*

One may wonder why our conditions for always peaceful equilibria do not depend on the costs of transgressions or hassling. This is due to our baseline assumption that t and h can only affect payoffs from negotiations, not from fighting. If the war payoffs were instead functions of these choices, then the players' reservation values for conflict would depend on the marginal effects and costs of transgressions and hassling. In order for the flexible-response framework to materially affect the prospects for peace, these responses must shape payoffs in war as well as peace.[23]

6.2 Terms of Settlement

In ordinary crisis bargaining games, the only way to get a better deal at the bargaining table is to run a greater risk of war (Banks, 1990, Lemma 3). As we observed in our analysis of private type and equilibrium payoffs above, flexible responses like hassling introduce another mechanism for states to obtain better terms from a settlement. Even with little or no threat of war, a state may use transgressions or hassling to shift the balance of bargaining power.

Because hassling is costly, any increase in hassling must come with a commensurate benefit in the terms of settlement. Otherwise, it would obviously be a profitable deviation if a state could obtain a better deal while incurring a lower cost of hassling. Consequently, whenever we compare two types that both end up at peace in the equilibrium of a flexible response crisis bargaining game, the one that hassles more must get a better deal. If both hassle the same amount, then they should receive identical settlements—just like states that run the same risk of war in ordinary crisis bargaining games.

[23] A similar mechanism drives results in Liu (2021).

Proposition 6. *If $\pi(\theta) = \pi(\theta') = 1$ and $h(\theta) \leq h(\theta')$, then $V_D(\theta) \leq V_D(\theta')$. Furthermore, if $h(\theta) < h(\theta')$, then $V_D(\theta) < V_D(\theta')$.*

Placing additional structure on the model primitives allows us to be even more specific about the relationship between the extent of hassling and the value of settlement. First, we will assume D's type is drawn from an interval, $\theta \in [\underline{\theta}, \bar{\theta}]$. This requirement effectively allows us to strengthen the incentive compatibility conditions for equilibrium, as we can now say that every type of D must find it unprofitable to mimic the strategy of a marginally stronger or weaker type. Second, we will assume a certain degree of differentiability (and thus continuity) in the relationship between private type and war payoffs, as well as that between private type, hassling amount, and the cost of hassling.[24] These assumptions allow us to characterize local incentive compatibility conditions—the lack of incentive to mimic a slightly lower or higher type—in terms of derivatives of the war payoff and hassling cost functions. We refer to the collection of these assumptions as bounded variation conditions, or (BV).

Definition 5. The model has *bounded variation* if W_D and K_D are differentiable and

$$\left.\begin{aligned} \Theta = [\underline{\theta}, \bar{\theta}] \qquad & \text{where } \underline{\theta} < \bar{\theta}, \\ |W_D(\theta) - W_D(\theta')| \leq M_W|\theta - \theta'| \qquad & \text{for all } \theta, \theta' \in \Theta, \text{ where } M_W < \infty, \\ |K_D(h, \theta) - K_D(h', \theta')| \leq \qquad & \text{for all } h, h' \in \mathcal{H} \\ M_D\|(h, \theta) - (h', \theta')\| \qquad & \quad \text{and } \theta, \theta' \in \Theta, \text{ where } M_D < \infty. \end{aligned}\right\} \tag{BV}$$

The bounded variation conditions allow us to apply the "envelope theorem" commonly employed in mechanism design analyses of crisis bargaining models (Banks, 1990; Fey and Ramsay, 2011). Given just a few endogenous elements of the equilibrium, we can determine every type's equilibrium payoff, which in turn will allow us to back out the precise terms of settlement for each type that ends up at peace. All we need to know are the lowest type's equilibrium utility,[25] whether each type ends up at war or peace, and the extent of hassling by those types that end up at peace. The following proposition gives a precise statement of $U_D(\theta)$ as a function of these equilibrium quantities.

[24] Specifically, we assume W_D and K_D are Lipschitz continuous, a weaker requirement than continuous differentiability.

[25] In fact, all that is necessary is to know the equilibrium payoff of a single type, not necessarily that of $\underline{\theta}$.

Lemma 4. *Assume* (BV) *holds. For all* $\theta_0 \in \Theta$,

$$U_D(\theta_0) = U_D(\underline{\theta}) + \int_{\underline{\theta}}^{\theta_0} (1 - \pi(\theta)) \frac{dW_D(\theta)}{d\theta}\, d\theta - \int_{\underline{\theta}}^{\theta_0} \pi(\theta) \left. \frac{\partial K_D(h, \theta)}{\partial \theta} \right|_{h=h(\theta)} d\theta. \tag{3}$$

This complex expression boils down to two essential facts about the relationship between private type and equilibrium payoffs. First, for types that go to war, the marginal increase in utility as θ increases is the same as the marginal increase in war payoff—naturally enough. Second, among those that settle in equilibrium, the marginal change in equilibrium payoff depends exactly on the marginal effect of private type on the cost of hassling. This second fact is what allows us to pin down the value of settlement once we know which types settle and how much they spend on hassling. Suppose there is an interval of types $[\theta', \theta''] \subseteq \Theta$ which all choose to settle in equilibrium: $\pi(\theta) = 1$ for all $\theta \in [\theta', \theta'']$. We can use Lemma 4 to characterize how the terms of the bargain differ between the poles of this interval:

$$V_D(\theta'') - V_D(\theta') = \underbrace{K_D(h(\theta''), \theta'') - K_D(h(\theta'), \theta')}_{\text{cost difference}} - \underbrace{\int_{\theta'}^{\theta''} \left. \frac{\partial K_D(h, \theta)}{\partial \theta} \right|_{h=h(\theta)} d\theta}_{\text{effectiveness premium}}.$$

The first term here, the cost difference, is a kind of baseline: incentive compatibility means the terms of settlement must adjust at least roughly in accordance with the cost paid for hassling. If not for the second term, the effectiveness premium, then each type $\theta \in [\theta', \theta'']$ would have the same equilibrium payoff. The effectiveness premium represents the additional benefit that states with greater hassling effectiveness can extract from bargaining. For example, suppose θ improves hassling capability, or that K_D is decreasing in θ. Then the effectiveness term will be positive, more so if there is a steep relationship between private type and the marginal cost of hassling. Conversely, there will be no benefit from the effectiveness premium if hassling does not take place. If $h(\theta) = 0$ for all $\theta \in [\theta', \theta'']$, then the cost difference and effectiveness premium are both zero, and all types in this interval receive the same settlement. In that case, we are back in the world of ordinary crisis bargaining games, where the only source of bargaining leverage is the threat of war.

In the special case of the bounded variation model where θ degrades hassling capability, we can even further pin down the value of settlement. We know from Proposition 1 that any equilibrium in this case will be characterized by a cutpoint $\hat{\theta} \in \Theta$, with all types below $\hat{\theta}$ settling in equilibrium and all types above it fighting. Using Lemma 4, we can then characterize the settlement value for all $\theta < \hat{\theta}$ in terms of the cutpoint type's war payoff and the choice of hassling by each intermediate type.

Corollary 2. *Assume θ degrades hassling effectiveness and (BV) holds. There exists $\hat{\theta} \in \Theta$ such that $\pi(\theta) = 1$ for all $\theta < \hat{\theta}$ and $\pi(\theta) = 0$ for all $\theta > \hat{\theta}$. If $\underline{\theta} < \hat{\theta} < \bar{\theta}$, then for all $\theta_0 < \hat{\theta}$,*

$$V_D(\theta_0) = W_D(\hat{\theta}) + K_D(h(\theta_0), \theta_0) + \int_{\theta_0}^{\hat{\theta}} \left. \frac{\partial K_D(h, \theta)}{\partial \theta} \right|_{h=h(\theta)} d\theta. \tag{4}$$

This result is useful for two reasons. First, it shows that we need relatively little information about the equilibrium to pin down settlement values in flexible response crisis bargaining games when θ degrades hassling capability. As long as we know the lowest type that fights and the amount of hassling exerted by each lower type, we can derive the exact settlement level in equilibrium. Importantly, if two bargaining games result in the same cutpoint type and choices of hassling below the cutpoint, they will also result in the exact same terms of settlement for each type of D, even if the bargaining processes themselves are quite dissimilar. Second, as we show in the next section, the necessary condition provided by Corollary 2 turns out to be sufficient. Specifically, for any non-decreasing hassling plan $h(\theta)$, if we allocate settlement values according to the given formula for V_D, the resulting mechanism is incentive compatible and satisfies voluntary agreements.

6.3 Amount of Hassling

We have shown that greater hassling effectiveness is associated with greater utility in case the crisis ends in a settlement (Proposition 3 and Proposition 4). We have also shown that the only way to improve one's payoff from a settlement is to choose greater levels of hassling (Proposition 6) Intuitively, then, it would appear to follow that more effective types hassle more in equilibrium. This intuition only holds in general when the cost of hassling satisfies the decreasing differences condition.

Proposition 7. *Assume (DD) holds. If $\pi(\theta) = \pi(\theta') = 1$ and θ' has greater hassling effectiveness than θ, then $h(\theta) \leq h(\theta')$.*

Because we can only say that more effective types hassle *weakly* more, this result leaves open two possibilities about why exactly more effective types have better equilibrium payoffs. One is that they hassle the same amount and receive the same settlement, so the higher payoff comes solely from the lower cost of hassling. The other possibility is that they hassle more and get better terms of settlement. Only if decreasing differences holds can we rule out a third possibility—that the more effective types hassle slightly less but at much lower cost, for a net increase in payoff despite the decrease in terms of settlement.

Combined with our earlier results on settlement values, Proposition 7 implies that the settlement value V_D increases with hassling effectiveness, as long as the decreasing differences condition holds. If θ enhances hassling effectiveness, this implies V_D is weakly increasing on the subset of Θ where peace prevails. If instead θ degrades hassling effectiveness, it means V_D is weakly increasing on $[\underline{\theta}, \hat{\theta})$, where $\hat{\theta}$ is the lowest type that weakly prefers to fight (see Corollary 2).

So far we have found that incentive compatibility implies that the level of hassling increases with effectiveness, as long as the decreasing differences condition holds. Can we say anything more specific about the relationship between private type and the equilibrium choice of hassling in a broad class of flexible response crisis bargaining games? If private strength is associated with lower hassling effectiveness, the answer turns out to be no—virtually any weakly decreasing hassling plan (subject to some continuity restrictions) can be sustained as the equilibrium of some bargaining game.

Proposition 8. *Assume θ degrades hassling effectiveness and* (BV) *and* (DD) *hold. Let h be any non-increasing and absolutely continuous function from $[\underline{\theta}, \bar{\theta}]$ into $\mathcal{H}$. Take any $\hat{\theta} \in (\underline{\theta}, \bar{\theta})$, let $\pi(\theta) = \mathbf{1}\{\theta \leq \hat{\theta}\}$, and let $V_D(\theta)$ be defined by Equation 4. The direct mechanism (h, π, V_D) satisfies* (IC) *and* (VA).

This result demonstrates that incentive compatibility and voluntary agreements alone place no restrictions on the pattern of hassling across types, besides that more effective types cannot engage in less of it. Consequently, the specifics of the relationship between effectiveness and the degree of hassling are highly model-dependent. For example, we cannot determine from the primitives alone whether all types hassle to the same degree, or whether there is some separation in levels of hassling. This will depend on how bargaining takes place and the precise effects of hassling choices on peacetime payoffs.

6.4 Limitations of the Flexible-Response Crisis Bargaining Framework

Our framework captures a crisis environment where states can form fully peaceful resolutions, states can conduct costly, low-level revisions to the peaceful resolution, or states can go to war. While the flexible-response crisis bargaining framework embraces a broader set of possible actions than what is allowed in standard crisis-bargaining models, it also has some subtleties and limits worth expanding on.

Our framework is distinct from what is covered in the Fey and Ramsay (2011) discussion of "war as a bargaining process" (pp. 166-167). In the setup in Fey and Ramsay or in the standard setup in models like Wagner (2000), Filson and Werner (2002), or Slantchev (2003a), there can be inefficiencies within bargained outcomes, which could be interpreted as allowing for low-level conflict to occur within bargaining. We concede that this interpretation is correct, but only if low-level conflict follows a very specific structure: namely, that the final payoffs engaging in some fighting and then settling are a convex-combination of the payoffs from fighting forever and the (efficient) payoff from the negotiated settlement. In the "war as a bargaining process" setup, each round of fighting has an identical structure, where the costs of ten rounds of fighting is identical to ten times the costs of one round of fighting. In the context of low-level conflict and war, this symmetry could be appropriate if low-level conflict used the same technology as a decisive war but was (proportionally) smaller or conducted over a shorter period. In most cases, this seems like too strong an assumption: sanctions, hassling, hybrid war, and gray zone conflict utilize different technologies and force postures than conventional war does, suggesting that these capabilities do not share a common per-round structure.

We highlight two shortcomings of our framework as avenues for future research. First, the flexible-response crisis bargaining framework does not speak to scenarios where transgressions or hassling decisions are noisily observed. This undercuts our framework's ability to describe, for example, settings where a transgression is imperfectly observed or where identifying attribution is hard. As one interpretation of these limitations, we admit that our framework is not well suited to describe cyberwarfare when attribution problems are present (Baliga *et al.*, 2020), or when there is a hidden development of technological capabilities as the transgression (Meirowitz *et al.*, 2008; Baliga and Sjöström, 2008; Schultz, 2010; Debs and Monteiro, 2014; Bas and Coe, 2016; Spaniel, 2019; Meirowitz *et al.*, 2019). Second, while introducing the possibility of a continuum of transgression and hassling options constitutes a step in the direction of better describing international interactions, our transgression and hassling options are both uni-dimensional. Building out a more sophisticated framework that allows for more dimensions of policy responses could better describe the world.

7 Conclusion

We examine a new class of models— flexible-response crisis bargaining models—using the tools of mechanism design. Flexible-response crisis bargaining models represent a useful adaptation of the standard, dichotomous crisis bargaining framework, where war and peace are the only possible outcomes; in our model states can engage in a continuum of conflict operations, which better captures international crises where states select from a continuum of possible conflict options (like gray zone conflict or sanctions). Rather than fix a single game form and solve for a set of equilibria, in our analysis, we identify the properties shared by all equilibria within the full class of flexible-response crisis bargaining games. This general analysis allows us to be confident that our results are not driven by specifics of the game form or a specific equilibria, but are ubiquitous to all flexible-response crisis bargaining models.

Our most surprising results are those that differ from the Banks (1990) and Fey and Ramsay (2011) monotonicity results. While existing research has shown that improved private war capabilities or an improved private willingness to go to war *always* results in weakly more war, this relationship is more nuanced when war capabilities can also benefit low-level conflict capabilities. Similarly, while Banks (1990) and Fey and Ramsay (2011) show that an improved private ability to conduct war *always* produces a greater utility, we find the results do not necessarily hold when a robust ability to go to war can hurt an actor's ability to effectively sanction or hassle.

A central concern of international relations is understanding the drivers of costly and destructive conflict. While this topic has been well examined through models of war and peace, much of what occurs in international relations falls outside of what could easily be classified as a peaceful bargain or a decisive war. While naturally any model must make some simplifying assumptions on how the world works, this paper shows that neglecting the possibility for low-level responses, we may be misunderstanding how what actually drives war. More work is needed on this topic.

References

Acharya, A., Grillo, E. *et al.* (2015) War with crazy types, *Political Science Research and Methods*, **3**, 281–307. (document)

Aghion, P. and Jackson, M. O. (2016) Inducing leaders to take risky decisions: dismissal, tenure, and term limits, *American Economic Journal: Microeconomics*, **8**, 1–38. (document)

Allen, S. H. and Martinez Machain, C. (2019) Understanding the impact of air power, *Conflict management and peace science*, **36**, 545–558. (document), 5

Ashworth, S. and Bueno de Mesquita, E. (2006) Monotone comparative statics for models of politics, *American Journal of Political Science*, **50**, 214–231. 2.1

Ashworth, S. and Sasso, G. (2019) Delegation to an overconfident expert, *The Journal of Politics*, **81**, 692–696. (document)

Axelrod, R. and Keohane, R. O. (1985) Achieving cooperation under anarchy: Strategies and institutions, *World politics*, **38**, 226–254. (document)

Baker, W. D. and Oneal, J. R. (2001) Patriotism or opinion leadership? the nature and origins of the "rally'round the flag" effect, *Journal of conflict resolution*, **45**, 661–687. (document), 5

Baliga, S., De Mesquita, E. B. and Wolitzky, A. (2020) Deterrence with imperfect attribution, *American Political Science Review*, **114**, 1155–1178. (document), 6.4

Baliga, S. and Sjöström, T. (2008) Strategic ambiguity and arms proliferation, *Journal of political Economy*, **116**, 1023–1057. 6.4

Baliga, S. and Sjöström, T. (2020) The strategy and technology of conflict, *Journal of Political Economy*, **128**, 3186–3219. 6

Banks, J. S. (1990) Equilibrium behavior in crisis bargaining games, *American Journal of Political Science*, pp. 599–614. (document), 1, 2.1, 11, 3, 4, 6.2, 6.2, 7

Baron, D. P. and Besanko, D. (1987) Commitment and fairness in a dynamic regulatory relationship, *The Review of Economic Studies*, **54**, 413–436. (document)

Baron, D. P. and Besanko, D. (1992) Information, control, and organizational structure, *Journal of Economics & Management Strategy*, **1**, 237–275. (document)

Bas, M. A. and Coe, A. J. (2016) A dynamic theory of nuclear proliferation and preventive war, *International Organization*, **70**, 655–685. 6.4

Benson, B. and Smith, B. C. (2020) Do alliances cause commitment problems and war?, *Working Paper*. 1

Berger, D. H., Gilday, M. M. and Schultz, K. L. (2020) *Advantage at Sea: Prevailing with Integrated All-Domain Naval Power*, United States Department of Defense. 5

Bolkcom, C. (2007) Military airlift: C-17 aircraft program, LIBRARY OF CONGRESS WASHINGTON DC CONGRESSIONAL RESEARCH SERVICE. 5

Buchanan, B. (2020) *The Hacker and the State: Cyber Attacks and the New Normal of Geopolitics*, Harvard University Press. 5

Chapman, T. L. and Reiter, D. (2004) The united nations security council and the rally'round the flag effect, *Journal of Conflict Resolution*, **48**, 886–909. (document), 5

Chassang, S. and Miquel, G. P. I. (2010) Conflict and deterrence under strategic risk, *The Quarterly Journal of Economics*, **125**, 1821–1858. 6

Chassang, S. and Padró i Miquel, G. (2009) Defensive weapons and defensive alliances, *American Economic Review*, **99**, 282–86. (document)

Cobzaş, Ş., Miculescu, R. and Nicolae, A. (2019) *Lipschitz Functions*, Springer. A.14

Coe, A. J. (2014) Sanctions as instruments of regime change: A case study of civil conflict in iraq during containment, Tech. rep., Working Paper. (document)

Coe, A. J. (2018) Containing rogues: A theory of asymmetric arming, *The Journal of Politics*, **80**, 000–000. (document), 1

Coe, A. J. and Vaynman, J. (2020) Why arms control is so rare, *American Political Science Review*, pp. 1–14. (document)

Debs, A. and Monteiro, N. P. (2014) Known unknowns: power shifts, uncertainty, and war, *International Organization*, **68**, 1–31. (document), 1, 6.4

Dell, M. and Querubin, P. (2018) Nation building through foreign intervention: Evidence from discontinuities in military strategies, *The Quarterly Journal of Economics*, **133**, 701–764. 5, 5

Di Lonardo, L. and Tyson, S. (2017) Political instability and the failure of deterrence, *Available at SSRN 3094929*. 6

Dunford, J. (2016) Gen. Dunford's Remarks and Q&A. (document)

Fearon, J. D. (1995) Rationalist explanations for war, *International organization*, **49**, 379–414. (document)

Fearon, J. D. (1996) Bargaining over objects that influence future bargaining power. 1

Fearon, J. D. (1997) Signaling foreign policy interests tying hands versus sinking costs, *Journal of Conflict Resolution*, **41**, 68–90. 1

Fey, M. and Kenkel, B. (2019) Is an ultimatum the last word on crisis bargaining? (document), 12

Fey, M. and Ramsay, K. W. (2009) Mechanism design goes to war: peaceful outcomes with interdependent and correlated types, *Review of Economic Design*, **13**, 233. (document), 16, 6.1

Fey, M. and Ramsay, K. W. (2011) Uncertainty and incentives in crisis bargaining: Game-free analysis of international conflict, *American Journal of Political Science*, **55**, 149–169. (document), 1, 2, 1, 2.1, 11, 15, 2.2, 2.2, 6.1, 6.1, 6.2, 6.4, 7

Filson, D. and Werner, S. (2002) A bargaining model of war and peace: Anticipating the onset, duration, and outcome of war, *American Journal of Political Science*, pp. 819–837. 6.4

Fuhrmann, M. and Kreps, S. E. (2010) Targeting nuclear programs in war and peace: A quantitative empirical analysis, 1941-2000, *Journal of Conflict Resolution*, **54**, 831–859. 1

Gannon, A. J., Gartzke, E., Lindsay, J. R. and Schram, P. (2020) The shadow of deterrence, *Manuscript*. (document), 1, 5

Gartzke, E. and Lindsay, J. R. (2015) Weaving tangled webs: offense, defense, and deception in cyberspace, *Security Studies*, **24**, 316–348. (document)

Gartzke, E. and Lindsay, J. R. (2017) Thermonuclear cyberwar, *Journal of cybersecurity*, **3**, 37–48. 1

Gurantz, R. and Hirsch, A. V. (2017) Fear, appeasement, and the effectiveness of deterrence, *The Journal of Politics*, **79**, 1041–1056. (document), 1

Harrison, T., Johnson, K., Roberts, T. G., Way, T. and Young, M. (2020) *Space threat assessment 2020*, Center for Strategic and International Studies (CSIS). 5

Hörner, J., Morelli, M. and Squintani, F. (2015) Mediation and peace, *The Review of Economic Studies*, **82**, 1483–1501. (document)

Horowitz, M. and Reiter, D. (2001) When does aerial bombing work? quantitative empirical tests, 1917-1999, *Journal of Conflict Resolution*, **45**, 147–173. (document), 5

Huth, P. K. (1999) Deterrence and international conflict: Empirical findings and theoretical debates, *Annual Review of Political Science*, **2**, 25–48. 6

Joseph, M. (2020) The secret failure of nuclear containment, *Working Paper*. (document)

Katz, Y. (2010) And they struck them with blindness, *The Jerusalem Post*. 1, 5

Kinzer, S. (2007) *Overthrow: America's century of regime change from Hawaii to Iraq*, Macmillan. 5

Kreps, S. E. and Fuhrmann, M. (2011) Attacking the atom: Does bombing nuclear facilities affect proliferation?, *Journal of Strategic Studies*, **34**, 161–187. (document), 5

Kydd, A. H. and McManus, R. W. (2017) Threats and assurances in crisis bargaining, *Journal of conflict resolution*, **61**, 325–348. 6

Lanoszka, A. (2016) Russian hybrid warfare and extended deterrence in eastern europe, *International Affairs*, **92**, 175–195. (document), 1

Levy, J. S. (2011) Preventive war: Concept and propositions, *International Interactions*, **37**, 87–96. 1

Liu, L. (2021) Domestic constraints in crisis bargaining, *Working paper*. (document), 23

Lyall, J. (2013) Airpower and coercion in counterinsurgency wars: Evidence from afghanistan, Tech. rep., Working Paper. 5, 5

Mazarr, M. J. (2015) Struggle in the gray zone, *War on the Rocks*. (document), 1

McCormack, D. and Pascoe, H. (2017) Sanctions and preventive war, *Journal of Conflict Resolution*, **61**, 1711–1739. (document), 1

Mearsheimer, J. J. *et al.* (2001) *The tragedy of great power politics*, WW Norton & Company. (document)

Meirowitz, A., Morelli, M., Ramsay, K. W. and Squintani, F. (2019) Dispute resolution institutions and strategic militarization, *Journal of Political Economy*, **127**, 378–418. 6.4

Meirowitz, A., Sartori, A. E. *et al.* (2008) Strategic uncertainty as a cause of war, *Quarterly Journal of Political Science*, **3**, 327–352. (document), 6.4

Meirowitz, A. *et al.* (2006) Designing institutions to aggregate preferences and information, *Quarterly Journal of political science*, **1**, 373–392. (document)

Milgrom, P. and Segal, I. (2002) Envelope theorems for arbitrary choice sets, *Econometrica*, **70**, 583–601. A.11

Myerson, R. B. (1979) Incentive compatibility and the bargaining problem, *Econometrica*, pp. 61–73. 2.2

Niou, E. M. and Ordeshook, P. C. (1990) Stability in anarchic international systems, *American Political Science Review*, **84**, 1207–1234. (document)

Opall-Rome, B. (2018) Declassified: How an israeli operation derailed syria's nuclear weapons drive, *Defense News*. 1

Owen, D. (2021) Time for a fully integrated dual-fleet force model, *War on the Rocks*. 5

Pape, R. A. (1996) *Bombing to win: Air power and coercion in war*, Cornell University Press. (document), 5

Paul, J. A. (2003) Oil companies in iraq: a century of rivalry and war, in *Global Policy Forum*, pp. 25–26. 5

Pietrucha, M. and Renken, J. (2019) Uplifted: The case for small tactical airlift, *War on the Rocks*. 5

Post, A. (2019) Flying to fail: Costly signals and air power in crisis bargaining, *Journal of Conflict Resolution*, **63**, 869–895. 4, 19

Powell, R. (1989) Nuclear deterrence and the strategy of limited retaliation, *The American Political Science Review*, pp. 503–519. (document)

Powell, R. (1999) *In the shadow of power: States and strategies in international politics*, Princeton University Press. (document)

Powell, R. (2006) War as a commitment problem, *International organization*, **60**, 169–203. 1

Powell, R. (2015) Nuclear brinkmanship, limited war, and military power, *International Organization*, **69**, 589–626. (document)

Ramsay, K. W. (2017) Information, uncertainty, and war, *Annual Review of Political Science*, **20**, 505–527. (document)

Reiter, D. (2005) Preventive attacks against nuclear programs and the "success" at osiraq, *Nonproliferation Review*, **12**, 355–371. 1

Schram, P. (2021) When capabilities backfire, *Manuscript*. (document), 5

Schultz, K. A. (1999) Do democratic institutions constrain or inform? contrasting two institutional perspectives on democracy and war, *International Organization*, **53**, 233–266. (document)

Schultz, K. A. (2010) The enforcement problem in coercive bargaining: Interstate conflict over rebel support in civil wars, *International Organization*, **64**, 281–312. (document), 1, 5, 6.4

Shachtman, N. (2011) The secret history of iraq's invisible war, *Wired, April*, **11**. 22

Slantchev, B. L. (2003a) The power to hurt: Costly conflict with completely informed states, *American Political Science Review*, **97**, 123–133. 6.4

Slantchev, B. L. (2003b) The principle of convergence in wartime negotiations, *American Political Science Review*, pp. 621–632. (document)

Slantchev, B. L. (2011) *Military threats: the costs of coercion and the price of peace*, Cambridge University Press. 2

Smith, A. (1998) International crises and domestic politics, *American Political Science Review*, **92**, 623–638. 6

Spaniel, W. (2019) *Bargaining over the bomb: the successes and failures of nuclear negotiations*, Cambridge University Press. 1, 6.4

Spaniel, W. and Bils, P. (2018) Slow to learn: Bargaining, uncertainty, and the calculus of conquest, *Journal of Conflict Resolution*, **62**, 774–796. (document)

Spaniel, W. and Malone, I. (2019) The uncertainty trade-off: Reexamining opportunity costs and war, *International Studies Quarterly*, **63**, 1025–1034. (document)

Stokes, D. (2007) Blood for oil? global capital, counter-insurgency and the dual logic of american energy security, *Review of International Studies*, pp. 245–264. 5

Wagner, R. H. (2000) Bargaining and war, *American Journal of Political Science*, pp. 469–484. (document), 6.4

Waltz, K. N. (1979) *Theory of international politics*, Waveland Press. (document)

Wendt, A. (1992) Anarchy is what states make of it: the social construction of power politics, *International Organization*, **46**, 391–425. (document)

Appendix

Contents

A Proofs

A.1 Proof of Lemma 1

This result, as well as Lemma 3 below, depends on the following auxiliary result on the expected utility from mimicking various types in the absence of hassling.

Lemma A.1. *If $h = 0$ and $\theta' < \theta''$,*

$$\Phi_D(\theta'' \,|\, \theta') \leq \Phi_D(\theta' \,|\, \theta') \leq \Phi_D(\theta' \,|\, \theta'') \leq \Phi_D(\theta'' \,|\, \theta'').$$

Proof. The first and third inequalities follow from (IC). The second follows because W_D is increasing and $h = 0$:

$$\begin{aligned}\Phi_D(\theta' \mid \theta') &= \pi(\theta')[V_D(\theta') - K_D(0,\theta')] + (1-\pi(\theta'))W_D(\theta') \\ &= \pi(\theta')V_D(\theta') + (1-\pi(\theta'))W_D(\theta') \\ &\leq \pi(\theta')V_D(\theta') + (1-\pi(\theta'))W_D(\theta'') \\ &= \pi(\theta')[V_D(\theta') - K_D(0,\theta'')] + (1-\pi(\theta'))W_D(\theta'') \\ &= \Phi_D(\theta'' \mid \theta').\end{aligned}$$

□

Lemma 1. *If $h = 0$ and $\theta' < \theta''$, then $\pi(\theta') \geq \pi(\theta'')$.*

Proof. Lemma A.1 implies

$$\Phi_D(\theta'' \mid \theta'') - \Phi_D(\theta'' \mid \theta') \geq \Phi_D(\theta' \mid \theta'') - \Phi_D(\theta' \mid \theta'),$$

which is equivalent to

$$(1-\pi(\theta''))[W_D(\theta'') - W_D(\theta')] \geq (1-\pi(\theta'))[W_D(\theta'') - W_D(\theta')].$$

As $W_D(\theta'') > W_D(\theta')$, this in turn implies $\pi(\theta') \geq \pi(\theta'')$. □

A.2 Proof of Proposition 1

Proposition 1. *Assume θ degrades hassling effectiveness. If $\theta' < \theta''$, then $\pi(\theta') \geq \pi(\theta'')$.*

Proof. For a proof by contradiction, suppose $\theta' < \theta''$ and $\pi(\theta') < \pi(\theta'')$. This implies $\pi(\theta') = 0$ and $\pi(\theta'') = 1$. (VA) implies

$$V_D(\theta'') - K_D(h(\theta''),\theta'') \geq W_D(\theta'').$$

(IC), combined with the assumption that θ' degrades hassling effectiveness, implies

$$W_D(\theta') \geq V_D(\theta'') - K_D(h(\theta''),\theta') \geq V_D(\theta'') - K_D(h(\theta''),\theta'').$$

Combining these inequalities gives $W_D(\theta') \geq W_D(\theta'')$, a contradiction. □

A.3 Proof of Proposition 2

Proposition 2. *Assume θ improves hassling effectiveness, and let $\theta' < \theta''$. If (WURI) holds, then $\pi(\theta') \geq \pi(\theta'')$. If (SURI) holds, then $\pi(\theta') \leq \pi(\theta'')$.*

Proof. We will prove the claims by contraposition. Let $\theta' < \theta''$, and suppose $\pi(\theta') < \pi(\theta'')$ (i.e., $\pi(\theta') = 0$ and $\pi(\theta'') = 1$). We want to prove that this implies (WURI) does not hold. Note that (VA) implies

$$V_D(\theta'') - K_D(h(\theta''), \theta'') \geq W_D(\theta''),$$

while (IC) implies

$$W_D(\theta') \geq V_D(\theta'') - K_D(h(\theta''), \theta').$$

Combining these gives

$$W_D(\theta') + K_D(h(\theta''), \theta') \geq V_D(\theta'') \geq W_D(\theta'') + K_D(h(\theta''), \theta''),$$

which in turn implies

$$W_D(\theta'') - W_D(\theta') \leq K_D(h(\theta''), \theta') - K_D(h(\theta''), \theta').$$

Because $\pi(\theta'') = 1$, this means (WURI) cannot hold, establishing the first claim of the lemma. An analogous argument establishes that (SURI) cannot hold if $\pi(\theta') > \pi(\theta'')$ (i.e., $\pi(\theta') = 1$ and $\pi(\theta'') = 0$). □

A.4 Proof of Lemma 2

Lemma 2. *Assume θ improves hassling effectiveness, (DD) holds, and $\max \mathcal{H} = \bar{h} < \infty$. If $W_D(\theta'') - W_D(\theta') > K_D(\bar{h}, \theta') - K_D(\bar{h}, \theta'')$ for all $\theta', \theta'' \in \theta'$ such that $\theta' < \theta''$, then (WURI) holds.*

Proof. For all $h < \bar{h}$ and $\theta' < \theta''$, (DD) implies

$$K_D(\bar{h}, \theta'') - K_D(h, \theta'') < K_D(\bar{h}, \theta') - K_D(h, \theta'),$$

which is equivalent to

$$K_D(\bar{h}, \theta') - K_D(\bar{h}, \theta'') > K_D(h, \theta') - K_D(h, \theta'').$$

Therefore, under the hypothesis of the lemma, we have

$$W_D(\theta'') - W_D(\theta') > K_D(\bar{h}, \theta') - K_D(\bar{h}, \theta'') \geq K_D(h, \theta') - K_D(h, \theta'')$$

for all $h \in \mathcal{H}$, which implies (WURI). □

A.5 Proof of Lemma 3

Lemma 3. *If $h = 0$ and $\theta' < \theta''$, then $U_D(\theta') \leq U_D(\theta'')$.*

Proof. Immediate from Lemma A.1. □

A.6 Proof of Proposition 3

Proposition 3. *Assume θ improves hassling effectiveness. If $\theta' < \theta''$, then $U_D(\theta') \leq U_D(\theta'')$. The inequality is strict if $\pi(\theta') = 0$ or $h(\theta') > 0$.*

Proof. When θ improves hassling effectiveness, (IC) implies

$$\begin{aligned}
U_D(\theta'') &= (1 - \pi(\theta''))W_D(\theta'') + \pi(\theta'')[V_D(\theta'') - K_D(h(\theta''), \theta'')] \\
&\geq (1 - \pi(\theta'))W_D(\theta'') + \pi(\theta')[V_D(\theta') - K_D(h(\theta'), \theta'')] \\
&\geq (1 - \pi(\theta'))W_D(\theta') + \pi(\theta')[V_D(\theta') - K_D(h(\theta'), \theta')] \\
&= U_D(\theta').
\end{aligned}$$

□

If $\pi(\theta') = 0$, then the second inequality above is strict. The same is true if $\pi(\theta') = 1$ and $h(\theta') > 0$.

A.7 Proof of Proposition 4

The proof depends on a more general property of hassling effectiveness and equilibrium utility:

Lemma A.2. *If $\pi(\theta') = \pi(\theta'') = 1$ and θ' has greater hassling effectiveness than θ'', then $U_D(\theta') \geq U_D(\theta'')$ (strictly if $h(\theta'') > 0$).*

Proof. By (IC), we have

$$U_D(\theta') \geq \underbrace{V_D(\theta'') - K_D(h(\theta''), \theta')}_{\Phi_D(\theta'' \mid \theta')} \geq V_D(\theta'') - K_D(h(\theta''), \theta'') = U_D(\theta''). \tag{A.1}$$

The second inequality of Equation A.1 is strict if $h(\theta'') > 0$. □

Proposition 4. *Assume θ degrades hassling effectiveness. There exists $\hat{\theta}$ such that $\pi(\theta) = 1$ for all $\theta < \hat{\theta}$ and $\pi(\theta) = 0$ for all $\theta > \hat{\theta}$. If $\theta' < \theta'' < \hat{\theta}$, then $U_D(\theta') \geq U_D(\theta'')$ (strictly if $h(\theta'') > 0$). If $\hat{\theta} < \theta' < \theta''$, then $U_D(\theta') < U_D(\theta'')$.*

Proof. The claim about $\hat{\theta}$ follows from Proposition 1. The next claim then follows from Lemma A.2. The final claim follows because W_D is strictly increasing in θ. □

A.8 Proof of Proposition 5

Proposition 5. *If $\hat{W}_C + W_D(\bar{\theta}) \leq 1$, then there is a flexible-response crisis bargaining game form with voluntary agreements that has an always peaceful equilibrium.*

Proof. We prove the result by construction. Consider the following extensive form game:

1. C chooses $t \in \mathcal{T}$.
2. D chooses $h \in \mathcal{H}$.
3. C and D simultaneously choose $b_C \in \{0, 1\}$ and $b_D \in \{0, 1\}$.

War occurs if either player chooses $b_i = 1$:

$$\pi^g(b_C, b_D) = \mathbf{1}\{b_C + b_D > 0\}.$$

Baseline payoffs are divided according to the war payoff for the strongest type of D, with each player paying a penalty $Z > 0$ if the other chooses a non-zero flexible response:

$$V_C^g(t, h, b_C, b_D) = 1 - W_D(\bar{\theta}) - Z \cdot \mathbf{1}\{h > 0\},$$
$$V_D^g(t, h, b_C, b_D) = W_D(\bar{\theta}) - Z \cdot \mathbf{1}\{t > 0\}.$$

Assume Z is arbitrarily large—specifically, $Z > \max\left\{1 - W_D(\bar{\theta}) - \hat{W}_C, W_D(\bar{\theta}) - W_D(\underline{\theta})\right\}$. Note that this game has voluntary agreements, as each player can guarantee war by choosing $b_i = 1$.

We claim that the following strategy profile constitutes an equilibrium of this game:

1. C chooses $t = 0$.
2. Following all choices of t, D chooses $h = 0$.
   - C's beliefs about D's type remain at the prior following all (t, h).
3. C chooses $b_C = 1$ if and only if $h > 0$. D chooses $b_D = 1$ if and only if $t > 0$.

Because Z was chosen to be arbitrarily large, the choices of bargaining strategy when $h > 0$ or $t > 0$ are clearly best responses. Now consider the case when $h = t = 0$. For C, deviating to $b_C = 1$ would result in an expected payoff of $\hat{W}_C \leq 1 - W_D(\bar{\theta})$, which is unprofitable. For any type of D, deviating to $b_D = 1$ would result in a payoff of $W_D(\theta) \leq W_D(\bar{\theta})$, which is unprofitable. The bargaining strategies therefore comprise a Bayesian Nash equilibrium. Moving up the game tree, a deviation by C to $t > 0$ or by D to $h > 0$ would result in war, which we have just shown is worse than the payoffs from the proposed strategies. Finally, note that C's beliefs are updated in accordance with Bayes' rule whenever possible. Therefore, the proposed strategy profile is a perfect Bayesian equilibrium. □

A.9 Proof of Corollary 1

Corollary 1. *If $W_C(\theta) = p - c_C$ and $W_D(\theta) = 1 - p - c_D(\theta)$, where $c_D : \Theta \to \mathbb{R}_+$, then there is a flexible-response crisis bargaining game form with voluntary agreements that has an always peaceful equilibrium.*

Proof. The result follows from Proposition 5, as $\hat{W}_C + W_D(\bar{\theta}) = 1 - c_C - c_D(\bar{\theta}) < 1$. □

A.10 Proof of Proposition 6

Proposition 6. *If $\pi(\theta) = \pi(\theta') = 1$ and $h(\theta) \leq h(\theta')$, then $V_D(\theta) \leq V_D(\theta')$. Furthermore, if $h(\theta) < h(\theta')$, then $V_D(\theta) < V_D(\theta')$.*

Proof. (IC) implies

$$V_D(\theta') - K_D(h(\theta'), \theta') \geq V_D(\theta) - K_D(h(\theta), \theta'),$$

which is equivalent to

$$V_D(\theta') - V_D(\theta) \geq K_D(h(\theta'), \theta') - K_D(h(\theta), \theta').$$

If $h(\theta) \leq h(\theta')$, then the RHS is non-negative, and the first claim follows. If $h(\theta) < h(\theta')$, then the RHS is strictly positive, and the second claim follows. $\square$

A.11 Proof of Lemma 4

We first state a helpful lemma.

Lemma A.3. *Assume (BV) holds. For all $\theta, \theta' \in \Theta$, Φ_D is differentiable with respect to θ, and*

$$\frac{\partial \Phi_D(\theta' \mid \theta)}{\partial \theta} = (1 - \pi(\theta'))\frac{dW_D(\theta)}{d\theta} - \pi(\theta')\frac{\partial K_D(h(\theta'), \theta)}{\partial \theta}.$$

Proof. The existence of $\partial \Phi_D / \partial \theta$ follows from (BV). The expression in the lemma then follows immediately from the definition of Φ_D. $\square$

We then rely on standard mechanism design arguments to establish the proposition.

Lemma 4. *Assume (BV) holds. For all $\theta_0 \in \Theta$,*

$$U_D(\theta_0) = U_D(\underline{\theta}) + \int_{\underline{\theta}}^{\theta_0} (1 - \pi(\theta))\frac{dW_D(\theta)}{d\theta}\, d\theta - \int_{\underline{\theta}}^{\theta_0} \pi(\theta) \left.\frac{\partial K_D(h, \theta)}{\partial \theta}\right|_{h=h(\theta)} d\theta. \tag{3}$$

Proof. (IC) implies $U_D(\theta) = \sup_{\theta' \in \Theta} \Phi_D(\theta' \mid \theta)$ for all $\theta \in \Theta$. Therefore, by Milgrom and Segal (2002, Theorem 1),

$$\frac{dU_D(\theta)}{d\theta} = \left.\frac{\partial \Phi_D(\theta' \mid \theta)}{\partial \theta}\right|_{\theta' = \theta}$$

at each point where U_D is differentiable. Furthermore, (BV) implies Φ_D is Lipschitz continuous in θ. The claim then follows from Lemma A.3 and Milgrom and Segal (2002, Corollary 1). $\square$

A.12 Proof of Corollary 2

Corollary 2. *Assume θ degrades hassling effectiveness and (BV) holds. There exists $\hat{\theta} \in \Theta$ such that $\pi(\theta) = 1$ for all $\theta < \hat{\theta}$ and $\pi(\theta) = 0$ for all $\theta > \hat{\theta}$. If $\underline{\theta} < \hat{\theta} < \bar{\theta}$, then for all $\theta_0 < \hat{\theta}$,*

$$V_D(\theta_0) = W_D(\hat{\theta}) + K_D(h(\theta_0), \theta_0) + \int_{\theta_0}^{\hat{\theta}} \left.\frac{\partial K_D(h, \theta)}{\partial \theta}\right|_{h=h(\theta)} d\theta. \tag{4}$$

Proof. Proposition 1 implies the existence of the cutpoint $\hat{\theta}$. We then have $U_D(\theta) = W_D(\theta)$ for all $\theta > \hat{\theta}$. (BV) implies that W_D is continuous and Lemma 4 implies that U_D is continuous, so if $\hat{\theta} \in (\underline{\theta}, \bar{\theta})$ we have $U_D(\hat{\theta}) = W_D(\hat{\theta})$. For $\theta_0 < \hat{\theta}$, Lemma 4 then gives

$$\begin{aligned} V_D(\theta_0) &= U_D(\theta_0) + K_D(h(\theta_0), \theta_0) \\ &= U_D(\hat{\theta}) + K_D(h(\theta_0), \theta_0) + \left[U_D(\theta_0) - U_D(\hat{\theta})\right] \\ &= W_D(\hat{\theta}) + K_D(h(\theta_0), \theta_0) + \int_{\theta_0}^{\hat{\theta}} \left.\frac{\partial K_D(h,\theta)}{\partial \theta}\right|_{h=h(\theta)} d\theta. \end{aligned}$$

□

A.13 Proof of Proposition 7

Proposition 7. *Assume (DD) holds. If $\pi(\theta) = \pi(\theta') = 1$ and θ' has greater hassling effectiveness than θ, then $h(\theta) \leq h(\theta')$.*

Proof. (IC) implies:

$$\begin{aligned} V_D(\theta) - K_D(h(\theta), \theta) &\geq V_D(\theta') - K_D(h(\theta'), \theta), \\ V_D(\theta') - K_D(h(\theta'), \theta') &\geq V_D(\theta) - K_D(h(\theta), \theta'). \end{aligned}$$

A rearrangement of terms gives

$$K_D(h(\theta'), \theta') - K_D(h(\theta), \theta') \leq V_D(\theta') - V_D(\theta) \leq K_D(h(\theta'), \theta) - K_D(h(\theta), \theta).$$

(DD) therefore implies $h(\theta) \leq h(\theta')$. □

A.14 Proof of Proposition 8

Proposition 8. *Assume θ degrades hassling effectiveness and (BV) and (DD) hold. Let h be any non-increasing and absolutely continuous function from $[\underline{\theta}, \bar{\theta}]$ into $\mathcal{H}$. Take any $\hat{\theta} \in (\underline{\theta}, \bar{\theta})$, let $\pi(\theta) = \mathbf{1}\{\theta \leq \hat{\theta}\}$, and let $V_D(\theta)$ be defined by Equation 4. The direct mechanism (h, π, V_D) satisfies (IC) and (VA).*

Proof. As a preliminary, note that because K_D is Lipschitz and h is absolutely continuous, $K_D(h(\theta), \theta)$ is absolutely continuous when viewed as a function of θ (Cobzaş *et al.*, 2019, Corollary 3.3.9). Consequently, V_D is absolutely continuous and thus differentiable almost everywhere on $[\underline{\theta}, \hat{\theta})$.

Now take any $\theta, \theta' \in \Theta$. If $\theta' < \hat{\theta}$, then

$$\begin{aligned}\Phi_D(\theta' \mid \theta) &= V_D(\theta') - K_D(h(\theta'), \theta)\\ &= W_D(\hat{\theta}) + K_D(h(\theta'), \theta') - K_D(h(\theta'), \theta) + \int_{\theta'}^{\hat{\theta}} \frac{\partial K_D(h, \theta'')}{\partial d\theta''}\bigg|_{h=h(\theta'')} d\theta''.\end{aligned}$$

Therefore, for almost all $\theta' < \hat{\theta}$, we have

$$\begin{aligned}\frac{\partial \Phi_D(\theta' \mid \theta)}{\partial \theta'} &= \frac{\partial K_D(h(\theta'), \theta')}{\partial h}\frac{dh(\theta')}{d\theta'} + \frac{\partial K_D(h(\theta'), \theta')}{\partial \theta'}\\ &\quad - \frac{\partial K_D(h(\theta'), \theta)}{\partial h}\frac{dh(\theta')}{\theta'} - \frac{\partial K_D(h(\theta'), \theta')}{\partial \theta'}\\ &= \underbrace{\frac{dh(\theta')}{\theta'}}_{\leq 0}\left[\frac{\partial K_D(h(\theta'), \theta')}{\partial h} - \frac{\partial K_D(h(\theta'), \theta)}{\partial h}\right].\end{aligned}$$

Because θ degrades hassling effectiveness, (DD) implies that the term in brackets is non-negative if $\theta \leq \theta'$ and non-positive if $\theta \geq \theta'$. Next, notice that

$$\begin{aligned}&\lim_{\theta' \to \hat{\theta}^-} \Phi_D(\theta' \mid \theta)\\ &\quad = \lim_{\theta' \to \hat{\theta}^-}\left[W_D(\hat{\theta}) + K_D(h(\theta'), \theta') - K_D(h(\theta'), \theta) + \int_{\theta'}^{\hat{\theta}} \frac{\partial K_D(h, \theta'')}{\partial d\theta''}\bigg|_{h=h(\theta'')} d\theta''\right]\\ &\quad = W_D(\hat{\theta}) + K_D(h(\hat{\theta}), \hat{\theta}) - K_D(h(\hat{\theta}), \theta).\end{aligned}$$

Therefore, if $\theta \leq \hat{\theta}$, then

$$\lim_{\theta' \to \hat{\theta}^-} \Phi_D(\theta' \mid \theta) \geq W_D(\hat{\theta}) \geq W_D(\theta) = \lim_{\theta' \to \hat{\theta}^+} \Phi_D(\theta' \mid \theta).$$

Conversely, if $\theta \geq \hat{\theta}$, then

$$\lim_{\theta' \to \hat{\theta}^-} \Phi_D(\theta' \mid \theta) \leq W_D(\hat{\theta}) \leq W_D(\theta) = \lim_{\theta' \to \hat{\theta}^+} \Phi_D(\theta' \mid \theta).$$

Finally, we have $\Phi_D(\theta' \mid \theta) = W_D(\theta)$ for all $\theta' > \hat{\theta}$. Altogether, these findings imply $\Phi_D(\theta' \mid \theta)$ is non-decreasing in θ' if $\theta' \in [\underline{\theta}, \theta]$ and non-increasing in θ' if $\theta' \in [\theta, \bar{\theta}]$. Therefore, (IC) holds. These results also imply $U_D(\theta) \geq \lim_{\theta \to \hat{\theta}^-} \Phi_D(\theta' \mid \theta) = W_D(\theta)$ for all $\theta \leq \hat{\theta}$, so (VA) also holds. $\square$