When Capabilities Backfire: How Improved Hassling Capabilities Produce Worse Outcomes

Peter Schram, Vanderbilt University

I formalize interactions between an endogenously rising challenger and a rival defender who can accept the challenger's rise, go to war before the rise comes to fruition, or degrade the challenger's growth through low-level conflict operations that I call "hassling." The novelty here is that the defender has private information about its hassling capabilities; this implies that the challenger does not know how much it can rise without provoking the defender to hassle or go to war. I find that when the defender's ability to hassle improves, it can provoke a strategic response in the challenger that undermines the defender's ability to use its private information productively and results in lower utility for the defender. This model provides insight into both Saddam Hussein's decision-making leading up to the 2003 US invasion and the stability-instability paradox.

Leading up to 2003, Saddam Hussein repeatedly denied entry to UN weapons inspectors and made false statements about Iraq's chemical and biological weapons. These behaviors led Iraq's adversaries to believe that Iraq was developing weapons of mass destruction (WMD) when, in reality, Iraq had no nuclear weapons, and its pursuit of WMD was limited. Saddam had good reason to keep inspectors out: he believed that the inspectors would detail the extent and location of his armaments, information that could be used maliciously by Iraq's adversaries (Coe and Vaynman 2020). Still, Saddam's behavior is puzzling given that it was the United States, a much more powerful state, that was demanding inspectors be allowed in. The 1990–91 Gulf War illustrated that the Iraqi army was no match for US forces. Additionally, transcripts from the Saddam regime suggest that Iraq knew that its behavior around WMD issues put it at risk for a confrontation with the United States. The well-known logic of commitment problems identifies this scenario as one in which preventive war is likely (Debs and Monteiro 2014; Fearon 1995; Gurantz and Hirsch 2017; Powell 2006), which is ultimately what happened. The puzzle, then, is why was Saddam Hussein not deterred by the might of the US military, when it seems that logically he should have allowed the inspectors in to avoid provoking a war he had little chance of winning.

Saddam kept weapons inspectors out in the lead up to 2003 because he was not only considering the Gulf War in his decision-making. In 1998, in response to Iraq turning away weapons inspectors, the United States conducted Operation Desert Fox, a four-day bombing campaign against Iraqi weapons facilities. In this operation, the United States illustrated that it had the ability and willingness to use targeted low-level conflict operations in response to rising powers. If, in the lead up to 2003, Saddam believed the United States was more likely to respond as it had in 1998 than it had in 1990, then there exists a troubling possibility: by being effective at low-level operations, the United States may have undermined its deterrent threat from war and emboldened Saddam to behave as he did. While low-level conflict capabilities can be useful in political crises—Operation Desert Fox was a success—they may also provoke problematic responses from opponents. Today, the availability of technologies such as cyberattacks, drone strikes, and precision strikes, tools that can cheaply and precisely destroy a nuclear program outside of a preventive war, may actually lead to more low-level conflict, or, in the case of the 2003 invasion, more war.

I formalize the above intuition. I consider a challenger who chooses how much to "transgress" (following Gurantz and Hirsch 2017), an act that will make the challenger more

Peter Schram (peter.schram@vanderbilt.edu) is an assistant professor at Vanderbilt University, Nashville, TN 37235.

Portions of this research were funded by the Center for Peace and Security Studies at University of California, San Diego, and the Office of Naval Research under award N000141912491. An appendix with supplementary material is available at https://doi.org/10.1086/719637.

Published online July 20, 2022.

powerful.[1] In response, a rival defender can accept the transgression, initiate a preventive war to stop the transgression, or "hassle," which, following Schram (2021), is the use of limited conflict to degrade the challenger's transgression while allowing bargaining to continue. What distinguishes this work from other formal research on conflict short of war (Bas and Coe 2016; Coe 2018; McCormack and Pascoe 2017; Schram 2021; Spaniel 2019) is that the defender's hassling capabilities have both public and private components. The challenger uses its observations of the defender's public hassling capabilities in making its transgression decision, but it does not know the defender's exact private willingness to hassle. Thus, the challenger does not know how much it can transgress before it provokes hassling or war. I show that when the defender possesses better public hassling capabilities, it can produce worse outcomes—more aggressive transgressions, more war, and more hassling—for the defender.

Improvements in hassling capabilities produce worse outcomes for the defender in two ways. First, improvements in the defender's hassling capabilities can *embolden the challenger*. When Iraq chose to turn away weapons inspectors and lie about its capabilities, it did so knowing that the United States might respond with force. Because Operation Desert Fox demonstrated that the United States was good at hassling, Iraq expected hassling (rather than war) to be the most likely response and determined that the benefits of turning away weapons inspectors outweighed the risks of the expected military response. In a counterfactual setting where the United States was worse at hassling, Iraq might have anticipated that refusing entry to inspectors would be met with a greater likelihood of war. In this counterfactual setting, the increased risk of war could have overshadowed whatever gains Iraq expected from turning away weapons inspectors and convinced Iraq to be more open. Because the United States was good at hassling, it may have incentivized risky behavior and emboldened Iraq to gamble.

Second, improvements in the defender's hassling capabilities can produce worse outcomes for the defender by *making the defender more predictable*. Possessing private information is useful because it allows for posturing to attain better outcomes. The threat of war is an effective deterrent; if a challenger thinks that a transgression will provoke its rivals to declare war, then the challenger will avoid that transgression. When the defender postures effectively, the defender convinces the challenger that it would go to war over a transgression when, in fact, it would not go to war over the transgression. Under some conditions, improvements in the defender's hassling capabilities diminish its ability to posture, thus allowing the challenger to calibrate its transgressions to extract more surplus from the defender. I examine this in the context of the proxy conflicts that occurred during the Cold War.

This article identifies new mechanisms for how improved low-level conflict capabilities can lead to worse outcomes for the actor making the improvements. Existing research has demonstrated similar results under alternate theoretical frameworks (Baliga, Bueno de Mesquita, and Wolitzky 2020; Bas and Coe 2016; Powell 1989, 2015; Schelling 1966; Spaniel and Malone 2019). For example, while Baliga et al. (2020) and Bas and Coe (2016) show that better low-level capabilities can undermine a deterrent threat through a "failure-to-burn-bridges" logic, those results would not arise if not for key model assumptions, namely, ambiguous attribution or a stochastically observed transgression process (respectively).[2] I show that improved low-level conflict capabilities can lead to worse outcomes when neither of the aforementioned assumptions are present, but there is uncertainty over an actor's willingness to conduct low-level conflict. This is not to say that existing results arise from empirically invalid assumptions—for example, Baliga et al.'s (2020) model is built to describe cyberattacks and terrorism with uncertain attribution. Rather, a strength of this article is that its key assumption follows from a foundational principle of security studies: that actors possess private information on their willingness to engage in conflict (Blainey 1988; Fearon 1995; Fey and Ramsay 2011; Ramsay 2017; Slantchev and Tarar 2011; Waltz 1979). This article's primary contribution is that it further clarifies the relationship between conflict capabilities and outcomes when there is a new type of private willingness to conduct conflict, thus advancing our understanding of how crises evolve when actors have multiple means of conducting conflict. Given the continued advancement of precision strike capabilities, information technologies, unmanned aerial vehicles, artificial intelligence, robotics, and other technologies—all of which can be used in low-level conflict and classified as hassling capabilities—this is a relationship worth thoroughly understanding.

While this article emphasizes variation in hassling capabilities, any factor shaping the information environment, like domestic political changes, espionage, or disclosures of private information, could produce "predictability" or "emboldening" effects. And, while this article empirically focuses on interactions between states, the model can apply to any interactions

1. Through the logic in Coe and Vaynman (2020) and Koblentz (2018) and the standard logic of endogenous military investments, Saddam's handling of weapons inspectors was a transgression, though, from the perspectives of Saddam and the United States, for different reasons.

2. See the appendix for a discussion on burning bridges as a distinct mechanism.

between actors engaging in multiple forms of conflict. Insurgent-insurgent violence (Schram 2019; Shapiro 2013) and violence by drug cartels (Cruz and Durán-Martínez 2016; Trejo and Ley 2021) are often targeted and restrained. Governments regularly engage in some kind of "hassling" against their citizens, whether as political repression (Ritter 2014) or extractive institutions (Acemoglu and Robinson 2012). Similarly, firms engage in a range of low-level conflict-type behaviors with rivals, including price wars, lawsuits, and predatory hiring.

THEORY

Terminology, scope, and interactions

Consider, as an example, Operation Outside the Box. In 2007, Israel discovered Syria was building a nuclear reactor. In the operation, Israel deployed an electronic warfare attack to disable Syrian air defenses and conducted an air strike to destroy the reactor (Katz 2010).

I will refer to actions like Syria developing a nuclear reactor as "transgressions." If transgressions are allowed to occur, then they strengthen the future conflict capabilities of the transgressor. Because the nuclear reactor moved Syria closer to being able to produce nuclear weapons, this was a transgression. Most simply, transgressions include endogenous investments in military conventional, nuclear, space, electronic, or cyber capabilities (Bas and Coe 2012, 2016; Buchanan 2020; Copeland 2001; Debs and Monteiro 2014; Lindsay and Gartzke 2018) or forming or supporting alliances (Benson and Smith 2020). Following the usage in Gurantz and Hirsch (2017), transgressions could also describe actions like securing valuable territory (Powell 2006), supporting militants in rival states (Schultz 2010), reneging on international security agreements (e.g., denying weapons inspectors access to facilities; Coe and Vaynman 2020), and others. Transgressions matter because through their transgressions, the challenger may create a commitment problem.

When a challenger undertakes a transgression and creates a commitment problem, rival actors may initiate a preventive war to stop the transgression (Bas and Coe 2016; Debs and Monteiro 2014). Instead of a this, Israel degraded Syria's nuclear facilities through an electronic warfare attack and air strike. I will refer to limited attacks designed to undercut the challenger's power shift as "hassling" (Schram 2021). The defender uses hassling to improve its future bargaining position when facing transgressions, where hassling functions as an alternative to accepting the challenger's power shift or fighting a decisive and costly war. Put another way, hassling operates like a steam valve: in an international system where a rising power may provoke rivals to declare a preventive war, hassling can diffuse the situation at a lower cost to both parties than war while allowing for negotiations to continue. As discussed in Schram (2021), limited strikes against nuclear facilities (Fuhrmann and Kreps 2010; Reiter 2005), hybrid conflict (Lanoszka 2016), gray-zone conflict (Gannon et al. 2020; Mazarr 2015), and (limited) preventive war (Levy 2011) could qualify as hassling. As characterized in this article, the term "hassling" could also apply to arming, sanctions, and containment regimes (Coe 2018; Joseph 2020; McCormack and Pascoe 2017).

This article examines the effects of improvements in "hassling capabilities," which are simply the tools that an actor uses to hassle. In the case of Operation Outside the Box, this was Israel's electronic warfare and air strike capabilities. While other work considers low-level conflict capabilities, the novelty here is that there is a private component to a defender's hassling capabilities and, thus, its willingness to hassle. Here, the challenger is uncertain over how much it can transgress before it provokes hassling or war.

I model hassling capabilities as having both public and private components to mirror real life. The public component can be interpreted as what other actors know for certain about the defender's hassling capabilities, which largely comes from military intelligence and observations of a defender's previous hassling operations. For example, in 2007, Syria knew that Israel had a robust hassling capability, as Syria had observed in Operation Opera (1981). The private component can be interpreted as what other actors do not know for certain about the defender's hassling capabilities, which largely comes from uncertainty over the defender's willingness to hassle or the defender's secret technological capabilities. Syria likely did not know that Israel possessed a weapon that could disrupt its air defenses (Katz 2010). That an actor possesses private information about its conflict capabilities is not new, and the claim that actors have private information about their hassling capabilities is analogous to claims that actors have private information about their war capabilities (see Ramsay [2017] for a review). Following this logic, a "public improvement in hassling capabilities" means improvements to the defender's ability to hassle that can be observed by rivals. These improvements could occur through publicly announced upgrades in weapons capabilities, training exercises for hassling-type operations, or strategic leaking of classified data on capabilities.

In the model, I do not assume that improvements in hassling capabilities also affect war payoffs, although I discuss this possibility in the "Additional Results" section. As some justification, there is an ongoing debate over the value of air strikes in war (Horowitz and Reiter 2001; Post 2019), with some case analyses suggesting that air strikes are more valuable in peacetime strikes against weapons facilities (i.e., in hassling) than in war (Kreps and Fuhrmann 2011). Thus, the capabilities for effective hassling may look quite different from the capabilities for an effective war, an observation that is given further

credence when noting how different the force postures of Operation Desert Fox (hassling) and the 2003 Iraq War (an interstate war then a counterinsurgency) were. In fact, there may even exist a negative relationship between hassling capabilities and war capabilities, especially when militaries face finite budgets and must make strategic allocation decisions.

I formalize the following interactions. First, the challenger chooses how aggressively to transgress. Upon observing the transgression, a rival defender can accept the transgression, go to war over the transgression, or hassle to degrade the transgression. Because the defender's response depends on its hassling capabilities, which have a private component, the challenger has uncertainty over how the defender will respond to the selected level of transgression. I consider how changing the defender's capabilities affects its final outcomes. The primary dynamic of interest is when an improvement in hassling capabilities creates overall worse outcomes for the defender—in other words, when there is no possible way for the defender to do better following the improvement in hassling capabilities.

Mechanisms

Here I describe how improved public hassling capabilities can make the defender more predictable or can embolden the challenger, thus leading to overall worse outcomes for the defender. This nontechnical discussion previews the more formal results below.

First, improvements in public hassling capabilities can make the defender more predictable. In the model, defenders exploit their private information about their hassling capabilities. War is an effective deterrent threat. A defender with private hassling capabilities can bluff or posture—acting as if they have inferior hassling capabilities—to convince the challenger that even minor transgressions would be met with war (and not hassling). If this posturing works, then the challenger selects less aggressive transgressions and the defender attains information rents. However, when the defender publicly becomes better at hassling, sometimes the challenger can better predict how the defender will behave. For example, because the United States had a robust ability to conduct low-level covert operations and because a war would be so costly, the Soviet Union likely knew that its invasion of Afghanistan in 1979 would not be met with the United States declaring war but could be met with hassling; thus, the Soviet Union did not need to scale back its transgression out of fear of war. When public improvements in hassling capabilities reduce the uncertainty about how the defender will respond, the defender cannot posture as effectively. Essentially, improvements in public hassling capabilities can diminish the defender's benefits from its private information.

Second, improvements in hassling technologies can embolden a challenger. When choosing the extent to transgress, the challenger faces a trade-off: a larger transgression is advantageous if the defender does not declare war, but a larger transgression increases the likelihood that the defender will declare war. The terms of this trade-off—whether a small increase in transgression extent will produce a small or large increase in the likelihood of war—are dictated by the defender's hassling capabilities, which have both public and private components. Under some trade-off terms—say an 8% increase in transgression level produces a 20% increase in the likelihood of war—the challenger is unwilling to increase its level of transgression. However, under different trade-off terms that could occur when the defender is better at hassling—a 27% increase in transgression level produces a 20% increase in the likelihood of war—the challenger would be willing to increase its level of transgression.[3] If the challenger is emboldened in this way, being better at hassling can actually cause worse outcomes for the defender to an extent that offsets its gains from being better at hassling.

If the defender is publicly better at hassling, it can result in a greater likelihood of war, more pervasive and intensive hassling campaigns, or both; in other words, improvements in hassling capabilities can produce overall worse outcomes for the defender. But improved public hassling capabilities do not always produce overall worse outcomes for the defender. A defender could improve its hassling capabilities by developing a wide range of complementary hassling tools that could be used separately or together. If this is the case, then improvements in hassling capabilities can improve the defender's outcomes by making the defender less predictable. Alternatively, if a defender becomes so effective at hassling that it could easily degrade a rival's transgression, then the challenger may choose not to pay the costs of transgressing. Of course, achieving this degree of hassling efficacy may be difficult, as rival countries may be willing to absorb huge costs if it means, for example, developing nuclear weapons. Furthermore, interestingly, I show that improvements in public hassling capabilities are distinct from improvements in private hassling or wartime capabilities, which cannot produce overall worse outcomes for the defender (as discussed below).

Related theory

This model is similar to Gurantz and Hirsch (2017) and other models with endogenous power shifts but differs in two key respects: it allows the defender to select from a continuum of low-level conflict options (i.e., hassling), and it assumes the challenger is uncertain over the defender's hassling capabilities.

3. These numbers follow from the parameters used in fig. 2.

This model is rooted in a tradition that examines endogenous transgressions (often arming) and deterrence, where a challenger undertakes some action that is detrimental to a defender, and a defender determines how best to respond. There are many variants of this interaction that consider uncertainty over the challenger's first move (Baliga and Sjöström 2008; Bas and Coe 2016; Debs and Monteiro 2014; Meirowitz and Sartori 2008; Meirowitz et al. 2019; Schultz 2010; Spaniel 2019), domestic political considerations (Di Lonardo and Tyson 2022; Fearon 1994; Smith 1998), deterrence over multiple threats (McMahon and Slantchev 2015; Paine 2021), signaling (Fearon 1997; Kydd and McManus 2017), uncertainty over a challenger's type (Acharya and Grillo 2015; Gurantz and Hirsch 2017; Trager 2010), a first mover's advantage (Baliga and Sjöström 2020; Chassang and Padró i Miquel 2010), ambiguous attribution (Baliga et al. 2020), and others (Huth 1999).

In nearly every model cited above, when an actor is dissatisfied with its outcome, it declares war. Of course, the real world is not this simple, and in practice, decision-makers select from many possible policies (Gannon et al. 2020; Gartzke and Lindsay 2017; Lanoszka 2016; Mehta and Whitlark 2017). An emerging branch of formal research considers a policy maker who faces a range of policy options, including reneging on bargained agreements (Schultz 2010), limited war (Powell 2015), or efforts that degrade a rising power (Coe 2018; Joseph 2020; McCormack and Pascoe 2017; Schram 2021). The existing literature has shown that low-level conflict can be an alternative to preventive war. This article builds on this finding but demonstrates that an enhanced ability to conduct low-level conflict can have negative strategic effects. In other words, while existing research shows an enhanced ability to conduct low-level conflict is useful within a crisis, this article shows it can also shape an opponent's behavior to produce riskier crises at higher frequencies.

The relationship between uncertainty and conflict is a central topic in contemporary international relations (Fearon 1995; Fey and Ramsay 2011; Morrow 1989; Slantchev and Tarar 2011; Wagner 2000), with Ramsay (2017) providing an excellent review. This article is the first to treat private conflict capabilities or willingness to fight as affecting different types of conflict differently, with this article focusing exclusively on uncertainty over an actor's ability and willingness to use hassling. The analysis conducted here is similar to related research that considers how changes in wartime capabilities (or the features of peacetime) affect outcomes like utilities and the likelihood of war (Benson, Meirowitz, and Ramsay 2016; Chassang and Padró i Miquel 2009; Spaniel 2020; Spaniel and Malone 2019).

Other research has also shown that improved low-level conflict capabilities can lead to worse outcomes, albeit under different assumptions and through different mechanisms. Powell (1989, 2015) assumes low-level conflict is a tool for imposing costs in a war of attrition or for generating an elevated risk of conflict (respectively). Baliga et al. (2020) and Bas and Coe (2016) assume transgressions are imperfectly attributable or stochastically observed (respectively). A particularly important (and uncommon) assumption made here is that the challenger can select from a flexible set of possible transgressions ($t \geq 0$) rather than a dichotomous "transgress or not" choice ($t \in \{0, 1\}$). This modeling choice rules out the mechanisms behind "burning bridges" (Schelling 1966, 1980) or "tying hands" (Fearon 1997) in the model below. With completely flexible transgressions and complete information, a challenger will always select a calibrated transgression designed to make the defender indifferent between going to war or not, thus always granting the defender its wartime utility regardless of its hassling capabilities. As I describe in corollary 3 and the appendix, improvements in hassling capabilities cannot produce overall worse outcomes for the defender unless private information also plays a role, and then only sometimes. As a result, the theoretical mechanisms (emboldening and predictability) are quite different from the seminal logic of burning bridges and more similar to research on information and crisis bargaining (Fearon 1995; Fey and Ramsay 2011; Meirowitz et al. 2019; Spaniel 2020).

MODEL

Two actors, C and D, are in a crisis over a divisible asset with a normalized value of 1. C wants to transgress (denoted by $t \geq 0$) to improve its future wartime capabilities. D observes the selected transgression, and D will accept the transgression, engage in hassling to degrade the transgression (denoted by $h \geq 0$), or go to war to prevent the transgression from coming to fruition. When D does not declare war, t and h affect C's future likelihood of winning in war, which affects future bargaining between C and D. Critically, D's costs from hassling are influenced by (nonnegative) public parameter $\alpha \in \mathcal{A} \subseteq \mathbb{R}_+$ and private type $\theta \in \Theta \subseteq \mathbb{R}_+$. The rationale for this is described above, but as one interpretation, α is D's observed hassling capacity, and θ is D's private willingness to hassle. In this setup, C does not know how much it can transgress before D will hassle or go to war. The game is as follows.

1. I define $\{\underline{\theta}, \bar{\theta}\} = \Theta$ with $\underline{\theta} < \bar{\theta}$. Nature sets $\theta = \underline{\theta}$ with probability $\Pr(\underline{\theta})$ and sets $\theta = \bar{\theta}$ with probability $1 - \Pr(\underline{\theta})$. D knows nature's selection θ, but C does not.
2. C selects transgression level $t \geq 0$.
3. D can either go to war by setting $w_D = 1$ or not go to war by setting $w_D = 0$ and selecting some level

Table 1. Summarized Payoffs for Actors

Scenario	C's Utility	D's Utility
D initiates war at stage 3 (before h and t are realized)	$P(0, 0) - \kappa_C$	$1 - P(0, 0) - \kappa_D$
C initiates war at stage 5 (after h and t are realized)	$P(t, h) - \kappa_C$	$1 - P(t, h) - \kappa_D - \frac{h^2}{F(\alpha,\theta)}$
C accepts at stage 5 (after h and t are realized)	x	$1 - x - \frac{h^2}{F(\alpha,\theta)}$

of hassling $h \geq 0$ (with $h = 0$ implying that D "accepts"). When D does not go to war, the game moves to the next stage. Going to war terminates the game and produces wartime payoffs $U_C = P(0,0) - \kappa_C$ and $U_D = 1 - P(0,0) - \kappa_D$ for actors C and D, respectively (I characterize the P function below).

4. D offers C some value $x \in [0, 1]$.
5. C can declare "war" by setting $w_C = 1$ or can "accept" the offer by setting $w_C = 0$. When C sets $w_C = 1$, C receives its updated wartime payoff $U_C = P(t,h) - \kappa_C$, and D receives $U_D = 1 - P(t,h) - \kappa_D - h^2/F(\alpha,\theta)$, which is D's updated wartime payoff $(1 - P(t,h) - \kappa_D)$ minus its costs from hassling $(h^2/F(\alpha,\theta))$. When C sets $w_C = 0$, C receives payoff $U_C = x$ and D receives $U_D = 1 - x - h^2/F(\alpha,\theta)$.

Stages 1–3 describe C's transgression and D's initial response. Stages 4 and 5 describe future crisis bargaining after the transgression and hassling have been realized. As I discuss in the appendix, similar results can arise through many different bargaining protocols outside of what is modeled in stages 4 and 5, or through similar modeling choices made in Gurantz and Hirsch (2017), which makes this bargaining a black box. As an intuition for the game form, in Operation Outside the Box, Syria's investment in the reactor would happen in stage 2, Israel's hassling would happen in stage 3, and future crisis bargaining between Syria and Israel after the operation occurred would happen in stages 4 and 5. Israel's decision to hassle in 3 was thus driven by its concern over how crisis bargaining with a nuclear-armed Syria would play out in 4 and 5. Note that the treatment of war as a game-ending move is consistent with the common treatment of war as a costly lottery (see table 1).

The function P is the likelihood that C wins in a war. I assume functional form $P(t,h) = \max\{\min\{1, \rho + t - h\}, \rho\}$ with $\rho \in [0,1]$, implying that P is weakly increasing in t and weakly decreasing in h and that P falls between ρ and 1 inclusive. The constant ρ is C's likelihood of winning a war before transgressions or hassling comes to fruition (when $t = 0$ and $h = 0$). The assumption that $P(t, h)$ must be weakly greater than ρ implies that hassling can undercut transgressions but hassling cannot make C a declining power, at most returning C to a baseline war victory likelihood of ρ.[4]

If D initiates war in stage 3, then C and D fight over the asset. C's likelihood of winning in war is $P(0,0) = \rho$, and $\kappa_C > 0$ and $\kappa_D > 0$ are C's and D's costs from war (respectively).

If C initiates war in stage 5, the fighting happens after the transgression and hassling have both occurred. The P function, κ_C, and κ_D have been discussed above. The expression $h^2/F(\alpha,\theta)$ denotes the additional costs that D incurs from hassling, where the function $F : \{\mathcal{A} \times \Theta\} \rightarrow \{\mathbb{R}_+ \setminus 0\}$, with F strictly increasing in α and θ. This functional form implies that D faces lower costs of hassling as α and θ increase and that D pays no costs when $h = 0$ (i.e., from accepting C's transgression). For ease, I assume that C faces no costs to transgress, which can be thought of as C facing low enough transgression costs such that these costs do not bind; I relax this assumption in the appendix, and the substantive results do not change.

If C accepts in stage 5, this represents a bargained outcome after the transgression and hassling have both occurred. The value x denotes the offer D makes to C.

EQUILIBRIUM CONCEPTS AND ASSUMPTIONS

I limit attention to pure strategy perfect Bayesian Nash equilibria. As I discuss in the appendix, considering mixed strategies does not change the substance of the results. I let σ denote a pair of strategies, or $\sigma = (\sigma_C, \sigma_D)$. I let $x^\star$, $h^\star$, $w_C^\star$, $w_D^\star$, $t^\star$, $\sigma_C^\star$, and $\sigma_D^\star$ denote equilibrium actions and strategies. Below I will abuse notation and use the term $\sigma^\star(\theta, \alpha)$ to represent equilibrium gameplay when D is type $\theta \in \{\underline{\theta}, \bar{\theta}\}$ and parameter $\alpha \in \{\underline{\alpha}, \bar{\alpha}\}$.

I also limit analysis below to scenarios in which the constraints on the P function and x do not bind. This implies that in equilibrium, $P(t^\star, h^\star) \in (\rho, 1)$ and $x^\star \in (0, 1)$. This assumption is useful because it eliminates the possibility that the kink in the P function or the requirement that $x \in [0, 1]$

4. While low-level conflict could potentially turn an actor into a declining power, this falls outside of the transgressing-hassling framework (and larger deterrence framework) that I adopt here. That said, the key results—that better hassling capabilities could lead to worse outcomes—would survive if hassling could make C decline but under different conditions.

drives any results, and it eliminates the need for excessive casework (details on this are in the appendix). In the appendix I include alternate functional forms and specifications that do not rely on this assumption, and the findings are substantively identical.

EQUILIBRIUM

The intuition is as follows. In the fourth and fifth stages, D does strictly better avoiding war, which D accomplishes by making C an offer equal to its wartime utility $x^\star = \rho + t^\star - h^\star - \kappa_C$. In the third stage, D reacts to C's selected transgression ($t^\star$) by either going to war ($w_D^\star = 1$) or not going to war ($w_D^\star = 0$) and selecting an optimal hassling level ($h^\star$). In the second stage, C selects a transgression $t^\star$ based on how C expects D will respond, which depends on D's known parameter (α) and expectations over D's private type (θ).

Because C faces no costs to transgress, C's utility from its transgression level is increasing unless it provokes D to war. Thus, C will attempt to avoid war with at least some types of D. Specifically, C will select a t that would make either a bad-at-hassling D ($D(\alpha, \underline{\theta})$) or a good-at-hassling D ($D(\alpha, \bar{\theta})$) indifferent to whether (a) D goes to war in stage 3 while C is weaker, or (b) D hassles and lets bargaining play out with a stronger-but-hassled C in stages 4 and 5. I let $t(\alpha, \theta)$ denote the level of transgression that would make a $D(\alpha, \theta)$ indifferent between war (implicitly in stage 3) and hassling (implicitly with the bargained outcome). If C selects $t(\alpha, \underline{\theta})$, C avoids war altogether. If C selects $t(\alpha, \bar{\theta})$, $D(\alpha, \underline{\theta})$ (but not $D(\alpha, \bar{\theta})$) will respond with war.

The equilibrium takes one of two cases: a case in which C avoids war and a case in which C sometimes risks war. Proposition 1 lists some features of the equilibrium, and the appendix contains the full equilibrium statement and all proofs.

Proposition 1: Under the assumptions above, for a fixed $\alpha \in \{\underline{\alpha}, \bar{\alpha}\}$, the following actions are part of the Perfect Bayesian Nash Equilibrium. Letting $Q(\alpha) = \Pr(\underline{\theta})(\kappa_C + \kappa_D) + [\Pr(\bar{\theta}) - \Pr(\underline{\theta})][F(\alpha, \underline{\theta})/4] - \Pr(\bar{\theta})[F(\alpha, \bar{\theta})/4]$:

Case 1. When $Q(\alpha) \geq 0$ holds, C selects transgression level $t^\star = t(\alpha, \underline{\theta})$, which results in both types of D hassling, setting $h^\star = F(\alpha, \theta)/2$ for all $\theta \in \{\underline{\theta}, \bar{\theta}\}$. $D(\alpha, \underline{\theta})$ attains $U_D(\sigma^\star(\underline{\theta}, \alpha)) = 1 - \rho - \kappa_D$, and $D(\alpha, \bar{\theta})$ attains utility $U_D(\sigma^\star(\bar{\theta}, \alpha)) = 1 - \rho - \kappa_D + \frac{F(\alpha, \bar{\theta})}{4} - \frac{F(\alpha, \underline{\theta})}{4}$.

Case 2. When $Q(\alpha) < 0$ holds, C selects transgression level $t^\star = t(\alpha, \bar{\theta})$, which results in $D(\alpha, \underline{\theta})$ declaring war and $D(\alpha, \bar{\theta})$ hassling, setting $h^\star = F(\alpha, \bar{\theta})/2$. Both $D(\alpha, \underline{\theta})$ and $D(\alpha, \bar{\theta})$ attain their wartime utility, $U_D(\sigma^\star(\theta, \alpha)) = 1 - \rho - \kappa_D$.

In case 1, C avoids war altogether by selecting a transgression level that makes the type of D that is least willing to hassle, type $\underline{\theta}$, indifferent between hassling and war, giving $D(\alpha, \underline{\theta})$ its wartime utility $1 - \rho - \kappa_D$. When C avoids war and selects $t(\alpha, \underline{\theta})$, the type of D that is more willing to hassle, type $\bar{\theta}$, will also hassle and will attain a utility above its wartime utility because hassling is cheaper for it. In case 2, C sometimes risks war by selecting a transgression level that makes the type of D that is most willing to hassle, type $\bar{\theta}$, indifferent between hassling and war, giving $D(\alpha, \bar{\theta})$ its wartime utility $1 - \rho - \kappa_D$. When C selects $t(\alpha, \bar{\theta})$, the type of D that is less willing to hassle, type $\underline{\theta}$, will want to go to war and will attain its wartime utility. Note that C attains different utilities across cases 1 and 2; $Q(\alpha)$ is simply the cut point where, for values below the cut point, C does better by selecting the higher $t^\star$—selecting $t(\alpha, \bar{\theta})$—and sometimes risking war relative to selecting $t(\alpha, \underline{\theta})$.

RESULTS

In this section, I discuss the outcome of interest: when improvements in hassling capabilities lead to D becoming more predictable or C becoming emboldened, thus producing overall worse outcomes for D. Formally, improvements in public hassling capabilities—moving from $\underline{\alpha}$ to $\bar{\alpha}$ with $\underline{\alpha} < \bar{\alpha}$—lead to overall worse outcomes when every type of D does worse following the improvements, or when $U_D(\sigma^\star(\theta, \underline{\alpha})) \geq U_D(\sigma^\star(\theta, \bar{\alpha}))$ for all $\theta \in \Theta$, and $U_D(\sigma^\star(\theta, \underline{\alpha})) > U_D(\sigma^\star(\theta, \bar{\alpha}))$ for at least some $\theta \in \Theta$.[5] I then discuss how these two mechanisms are the *only* way for D to experience overall worse outcomes and extensions.

Predictability

In the game, D benefits from its private information. An improvement in public hassling capabilities can result in different private types of D playing the game in a more similar manner. For example, if it is common knowledge that D is outstanding at cyberwarfare, then C may be better able to predict how D will respond to a wide range of selected transgressions (e.g., with cyberattacks). As a result, when public capabilities are higher, C can select a transgression that better extracts the benefit that D would otherwise attain from its private information. This can occur despite the likelihood of

5. I discuss why this definition is preferred to using D's ex ante expected utility (a less strict definition) in the appendix. Also, I include more details on these results in the appendix.

war remaining unchanged. Figure 1 visualizes this intuition. For the selected parameters, under both $\underline{\alpha}$ and $\bar{\alpha}$, the equilibrium is described in case 1 in proposition 1, in which C always avoids war.

Because C observes D's parameter α but not D's type θ, C selects $t(\alpha, \underline{\theta})$ to avoid war. In the top two panels of figure 1 (where $\alpha = \underline{\alpha}$), C's optimal transgression level $t^\star = t(\underline{\alpha}, \underline{\theta})$ is indicated by the asterisks. In the top panel, the selected $t^\star$ makes $D(\underline{\alpha}, \underline{\theta})$ indifferent between war and hassling, implying that $D(\underline{\alpha}, \underline{\theta})$ will attain its wartime utility. In the second panel, the space between the selected transgression and the transgression level that would have made a $D(\underline{\alpha}, \bar{\theta})$ indifferent between hassling and going to war (marked with a dashed line) represents D attaining some surplus. Instead of always receiving its wartime payoff, type $\bar{\theta}$ is better off *due to its private information*. In other words, so long as $D(\underline{\alpha}, \bar{\theta})$ keeps its type private in the lead-up to C's selection of t, $D(\underline{\alpha}, \bar{\theta})$ can attain some bargaining surplus.

However, when D's public hassling capabilities improve, D's private information may become less valuable. Because C has some uncertainty over how willing D is to hassle, C must benchmark its transgression against type $\underline{\theta}$ if C seeks to avoid war. If improvements in D's known hassling abilities reduce the importance of D's unknown type, then C does not need to scale back as much to prevent war. Comparing the top two and bottom two panels illustrates how this may occur. In the top two panels, for $\underline{\alpha}$, there is a significant gap between C's selected transgression $t(\underline{\alpha}, \underline{\theta})$ and the point that makes $D(\underline{\alpha}, \bar{\theta})$ indifferent between war and hassling ($t(\underline{\alpha}, \bar{\theta})$). Moving to $\bar{\alpha}$, there is a much smaller gap between $t(\bar{\alpha}, \underline{\theta})$ and $t(\bar{\alpha}, \bar{\theta})$. This implies that as D becomes publicly better at hassling (moving from $\underline{\alpha}$ to $\bar{\alpha}$), C is able to select a $t^\star$ that is closer to the point that gives $D(\bar{\alpha}, \bar{\theta})$ its wartime payoff. Put another way, in the bottom two panels it does not matter much if D's private type is revealed because $D(\bar{\alpha}, \bar{\theta})$ does not attain much surplus from its private information; D's public improvements in hassling have made D's private type a smaller factor in determining D's selected hassling levels and have therefore made D more predictable. This makes all types of D worse off because an increase in hassling capabilities means C can select better tailored transgressions. Proposition 2 defines the conditions under which an improvement in public hassling capabilities makes D more predictable, thus leading to overall worse outcomes for D.

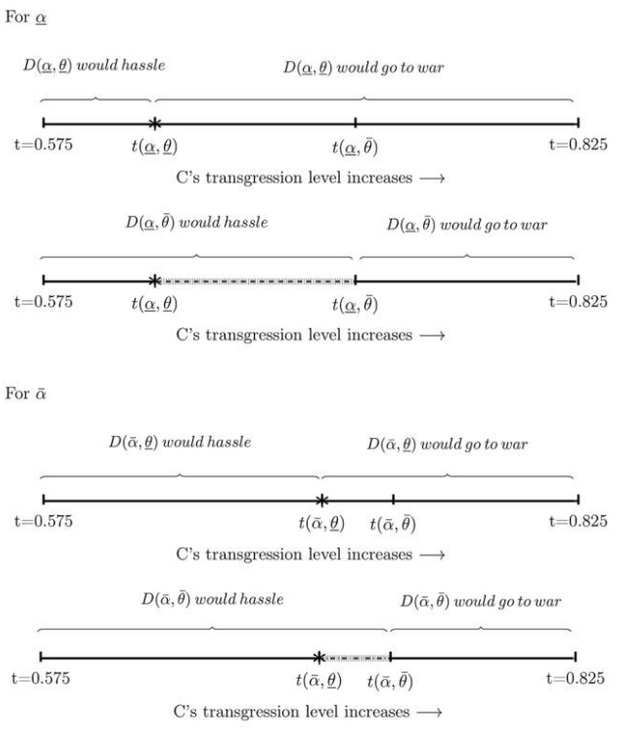

Figure 1. Optimal transgression and D's response (predictability). C's selected level of transgression under parameters $\underline{\alpha}$ and $\bar{\alpha}$ are denoted by the asterisks. D's response to the selected t's are bracketed off. Dashed lines represent the surplus $D(\alpha, \bar{\theta})$ attains from its private information in equilibrium. Parameters are $\kappa_D = 0.1$, $\kappa_C = 0.4$, $\rho = 0.5$, $\Pr(\underline{\theta}) = 0.5$, $\Pr(\bar{\theta}) = 0.5$, $F(\underline{\alpha}, \underline{\theta}) = 0.5$, $F(\underline{\alpha}, \bar{\theta}) = 1$, $F(\bar{\alpha}, \underline{\theta}) = 0.8$, and $F(\bar{\alpha}, \bar{\theta}) = 1.1$.

Proposition 2 (Predictability): Under the predictability conditions, C avoids war across parameters $\underline{\alpha}$ and $\bar{\alpha}$; formally, $Q(\underline{\alpha}) \geq 0$ (condition 1) and $Q(\bar{\alpha}) \geq 0$ (condition 2). Additionally, D's private information plays a diminished role under parameter $\bar{\alpha}$ relative to parameter $\underline{\alpha}$; formally, $F(\underline{\alpha}, \bar{\theta}) - F(\underline{\alpha}, \underline{\theta}) > F(\bar{\alpha}, \bar{\theta}) - F(\bar{\alpha}, \underline{\theta})$ (condition 3). When these conditions hold, improvements from $\underline{\alpha}$ to $\bar{\alpha}$ lead to overall worse outcomes for D.

Proposition 2 generalizes the intuition in figure 1. The first two predictability conditions imply that across $\underline{\alpha}$ and $\bar{\alpha}$, the equilibrium is characterized in case 1 of proposition 1. The third condition, which is decreasing differences in F (Ashworth and Bueno de Mesquita 2006), captures the differences in how types $\underline{\theta}$ and $\bar{\theta}$ play the game across parameters $\underline{\alpha}$ and $\bar{\alpha}$. Intuitively, across α's, C selects $t(\alpha, \underline{\theta})$ to make $D(\alpha, \underline{\theta})$ indifferent between war and hassling. Because $D(\alpha, \bar{\theta})$ faces lower costs from hassling, $D(\alpha, \bar{\theta})$ selects a greater level of hassling, captured in the optimal hassling level $h^\star(\theta) = F(\alpha, \theta)/2$, and attains a greater utility, captured in the $[F(\alpha, \bar{\theta})/4] - [F(\alpha, \underline{\theta})/4]$ term in $D(\alpha, \bar{\theta})$'s utility function. When $F(\alpha, \bar{\theta}) - F(\alpha, \underline{\theta})$ is small, $D(\alpha, \underline{\theta})$ and $D(\alpha, \bar{\theta})$ play the game similarly, implying that D's private information does not give $D(\alpha, \bar{\theta})$ much surplus. When the three conditions hold, the improvement in α degrades the value of D's private information, thus producing the outcome of interest.

Emboldening

In the game, C wants to transgress as much as possible but may be deterred from transgressing if further transgressions lead to a greater probability of war. Whether C is willing to transgress to a level that risks war is contingent on the benefit that C receives from the transgression, which is dependent on D's willingness to hassle. If C is emboldened, D's shift in hassling capabilities incentivizes C to transgress more aggressively, knowing that, while with some probability D may initiate war, C holds a better bargaining position when D does not initiate war. An improvement in public hassling capabilities emboldens C when the benefit of transgressing more aggressively outweighs the greater risk of war.[6] Figure 2 visualizes this intuition. C's expected utility is increasing in the selected t (moving right along the X-axis), until it provokes $D(\alpha, \underline{\theta})$ to go to war at $t(\alpha, \underline{\theta})$ (producing the first discontinuity), and is increasing in t again until all types to go to war at $t(\alpha, \bar{\theta})$.

The top panel of figure 2 describes the game under parameter $\underline{\alpha}$, where the gameplay is described in case 1 in proposition 1. In this game, C optimally selects $t^{\star} = t(\underline{\alpha}, \underline{\theta})$, and both $D(\underline{\alpha}, \underline{\theta})$ and $D(\underline{\alpha}, \bar{\theta})$ will always hassle. It is worth highlighting that C is not selecting $t(\underline{\alpha}, \bar{\theta})$. This "move not taken" captures C's trade-off between a greater transgression and a higher risk of war. The upside to selecting $t(\underline{\alpha}, \bar{\theta})$ is that C would have an advantage when facing $D(\underline{\alpha}, \bar{\theta})$.[7] However, the downside for C of selecting $t(\underline{\alpha}, \bar{\theta})$ is that now $D(\underline{\alpha}, \underline{\theta})$ will go to war, increasing the likelihood of war from zero to $\Pr(\underline{\theta})$. The difference between $t(\underline{\alpha}, \underline{\theta})$ and $t(\underline{\alpha}, \bar{\theta})$ can be thought of as the potential upside if C were willing to go to war with likelihood $\Pr(\underline{\theta})$. When $\alpha = \underline{\alpha}$, C's upside to doing better against $D(\underline{\alpha}, \bar{\theta})$ does not outweigh the costs of going to war with $D(\underline{\alpha}, \underline{\theta})$. In figure 2, this is shown by the asterisk at $t = t(\underline{\alpha}, \underline{\theta})$ giving C greater utility than the point at $t = t(\underline{\alpha}, \bar{\theta})$.

When $\alpha = \bar{\alpha}$, C faces a new trade-off, and the gameplay is described in case 2 in proposition 1. Selecting $t(\bar{\alpha}, \bar{\theta})$ will provoke $D(\bar{\alpha}, \underline{\theta})$ to go to war. However, as the bottom panel of figure 2 illustrates, C now faces a greater upside to selecting $t(\bar{\alpha}, \bar{\theta})$ over $t(\bar{\alpha}, \underline{\theta})$ because these cut points are so different, representing a large opening for a greater transgression if C is willing to risk some likelihood of war.[8] Now, C's upside to doing better against $D(\bar{\alpha}, \bar{\theta})$ outweighs the increased likelihood of war, as shown by the asterisk at $t(\bar{\alpha}, \bar{\theta})$ giving C greater utility than $t(\bar{\alpha}, \underline{\theta})$. This case leads to overall worse outcomes for D because an increase in hassling capabilities emboldened C to pursue a riskier transgression strategy and sometimes provoke war, both to D's detriment. Proposition 3 defines the conditions for improvements in public hassling capabilities to embolden C and lead to overall worse outcomes for D.

> **Proposition 3** (Emboldening): Under the emboldening conditions, C avoids war under parameter $\underline{\alpha}$ and goes to war under parameter $\bar{\alpha}$; formally, $Q(\underline{\alpha}) \geq 0$ (condition 1) and $Q(\bar{\alpha}) < 0$ (condition 2). When these emboldening conditions hold, improvements from $\underline{\alpha}$ to $\bar{\alpha}$ produce overall worse outcomes for D.

Proposition 3 generalizes the intuition in the example above. The $Q(\underline{\alpha}) \geq 0$ condition implies that C will not risk war under parameter $\underline{\alpha}$, which will give $D(\underline{\alpha}, \underline{\theta})$ its wartime utility and $D(\underline{\alpha}, \bar{\theta})$ a utility above its wartime payoff. The $Q(\bar{\alpha}) < 0$ condition implies that C will risk war under parameter $\bar{\alpha}$, which will result in both $D(\bar{\alpha}, \underline{\theta})$ and $D(\bar{\alpha}, \bar{\theta})$ attaining their wartime utilities. The outcome of interest arises because $D(\bar{\alpha}, \bar{\theta})$ does strictly worse than $D(\underline{\alpha}, \bar{\theta})$.

When (and only when) improved capabilities produce worse outcomes

The aim of this article is to define the relationship between capabilities and outcomes when there is uncertainty over the willingness to use low-level conflict. I showed that an improvement in α when the predictability or emboldening conditions hold represents two ways D can experience overall worse outcomes. But are there other ways for D to experience overall worse outcomes? Proposition 4 and the corollaries that follow show that there are not.

> **Proposition 4:** Improvements in α produce overall worse outcomes for D if and only if the emboldening or predictability conditions hold.

Proposition 4 shows that the two mechanisms characterized above represent all the ways that improvements in public hassling capabilities can produce overall worse outcomes for D. But can similar outcomes arise by considering private hassling capabilities or war capabilities?

> **Corollary 1:** Improvements in private hassling capabilities θ (e.g., moving from $\underline{\theta}$ to $\bar{\theta}$) cannot produce overall worse outcomes for D.

6. Spaniel and Malone (2019) find that improved economic interdependence can produce a similar result.

7. For the selected parameters, when nature sets $\theta = \bar{\theta}$, if C selects $t = t(\underline{\alpha}, \underline{\theta}) = 0.625$, C attains utility 0.375, and if C selects $t = t(\underline{\alpha}, \bar{\theta}) = 0.675$, C attains utility 0.425.

8. For the selected parameters, when nature sets $\theta = \bar{\theta}$, if C selects $t = t(\bar{\alpha}, \underline{\theta}) = 0.65$, C attains utility 0.1, and if C selects $t = t(\bar{\alpha}, \bar{\theta}) = 0.825$, C attains utility 0.275.

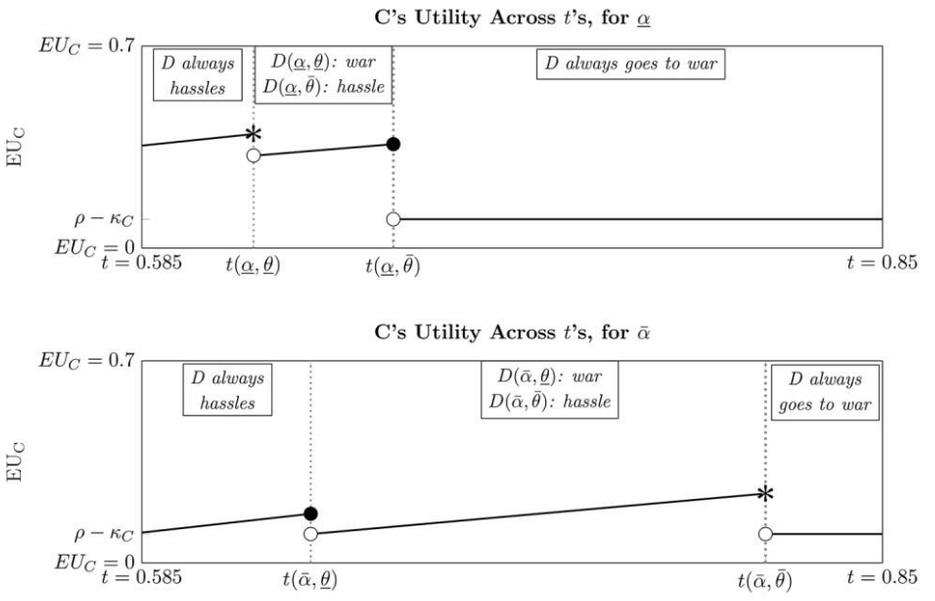

Figure 2. Optimal transgression and C's utility (emboldening). C's selected level of transgression under parameters $\underline{\alpha}$ and $\bar{\alpha}$ are denoted by the asterisks. D's response to the selected t's are labeled. Parameters are $\kappa_D = 0.1$, $\kappa_C = 0.4$, $\rho = 0.5$, $\Pr(\underline{\theta}) = 0.2$, $\Pr(\bar{\theta}) = 0.8$, $F(\underline{\alpha}, \underline{\theta}) = 0.5$, $F(\underline{\alpha}, \bar{\theta}) = 0.7$, $F(\bar{\alpha}, \underline{\theta}) = 0.6$, and $F(\bar{\alpha}, \bar{\theta}) = 1.3$.

Corollary 2: Improvements in war capabilities (e.g., decreasing ρ or κ_D) cannot produce overall worse outcomes for D.

Both corollaries follow proposition 1 and suggest that there is something unique about public hassling capabilities. And, analysis in the appendix demonstrates that these results are not simply theoretical anecdotes driven by a game form but rather can apply to a broad class of games and bargaining protocols with discrete or continuous type spaces. On corollary 1, because in response to a fixed t^* a privately-good-at-hassling D could always mimic the actions of privately-worse-at-hassling D and do weakly better, private improvements in hassling capabilities cannot generate overall worse outcomes for D.[9] On corollary 2, in any game in which some types attain their wartime utility, improvements in wartime payoffs always produce better outcomes for some types. Of course, it would be incorrect to say that public hassling capabilities are alone responsible for D facing overall worse outcomes.

Corollary 3: Improvements in public hassling capabilities cannot produce overall worse outcomes for D without private information (e.g., when $\underline{\theta} = \bar{\theta}$).

Corollary 3 follows from propositions 1 and 4. When $\underline{\theta} = \bar{\theta}$, C optimally selects a transgression level that makes D indifferent between war and hassling, guaranteeing D its wartime utility. In every iteration of the model that is considered here and in the appendix, interactions between public and private capabilities following improvements in public capabilities are needed for D to incur overall worse outcomes.

Additional results

What if hassling capabilities also affect wartime capabilities? Following the logic from corollary 2, if improvements in public hassling capabilities also improve payoffs from war, then the improvements in public hassling capabilities cannot produce overall worse outcomes for D. Because type $\underline{\theta}$ D always attains its wartime payoff, if a shift from $\underline{\alpha}$ to $\bar{\alpha}$ resulted in D attaining a greater wartime payoff, then $D(\bar{\alpha}, \underline{\theta})$ would always do better than $D(\underline{\alpha}, \underline{\theta})$. As discussed above, whether improved hassling capabilities affect war capabilities is an empirical question grounded in how different war operations and hassling operations are. At a minimum, this observation suggests a useful heuristic for developing hassling capabilities: so long as improvements in hassling also improve wartime outcomes, then these improvements will not produce overall worse outcomes for D.

What is the relationship between hassling capabilities and the likelihood of war? While not the main focus of the article, the relationship between wartime capabilities and the likelihood of war has received considerable attention (Benson et al. 2016; Kydd 2000; Powell 1999; Spaniel 2020). Similar to recent findings on war capabilities (see Benson et al. 2016),

9. For the broader class of games, this logic follows from standard mechanism design results where higher types attain an increasing information rent (Myerson 1979; Salanié 2005).

I find that improvements in public hassling capabilities could lead to more, less, or no change in the final likelihood of war. Thus, the relationship between low-level capabilities and the outbreak of war is nuanced, suggesting that, for example, third-party assistance to a protege's hassling capabilities could be stabilizing or destabilizing.

What if transgressions come with costs? I examine a model with transgression costs in the appendix, showing that while the results remain substantively similar, this alternate model imposes new, stricter predictability and emboldening conditions. These new conditions are needed because improving public hassling capabilities can produce a new effect: it can deter a challenger from transgressing in the first place. Whenever improving public hassling capabilities produces this effect, then all types of D cannot be weakly worse off.

EXAMINATION OF CASES

I now describe the theoretical mechanisms within the context of the 2003 US invasion of Iraq and of low-level operations during the Cold War. What follows should not be viewed as a causally identified test of the theory but rather as illustrations of the mechanisms.

Emboldening: Saddam's gamble

Within the cease-fire terms of the 1991 Persian Gulf War and following UN Security Council resolutions, Iraq was forbidden from possessing nuclear, biological, or chemical weapons (WMDs) and was subject to inspections from a UN special commission and the International Atomic Energy Agency. Between 1991 and 1998, Iraq repeatedly violated the terms of the resolutions by engaging in the following actions: preventing inspectors from searching suspected vehicles (1991), preventing inspectors from searching government buildings (1992), submitting incomplete disclosures of weapons programs (1995–96), and terminating all cooperation with inspectors (1998). These actions (and others) led US and British forces to conduct Operation Desert Fox, a bombing campaign against Iraqi military targets in 1998. Between 1999 and 2002, Iraq continued keeping weapons inspectors out, despite intensifying rhetoric within the United States. And, while in late 2002 Iraq agreed to resume inspections, after over a decade of friction with inspectors, Iraq's change proved insufficient, and the United States invaded Iraq in March 2003. In 2004 the Iraq Survey Group released the "Duelfer Report," which described how Iraq's WMD programs had essentially been destroyed in 1991 (Duelfer 2004).

To claim that Iraq was emboldened in the lead up to the 2003 US invasion, three questions must be addressed: Was Iraq "transgressing" as defined above? Did the United States demonstrate a robust hassling capability? Finally, did US hassling capabilities shape Iraqi behavior?

Saddam's transgression was keeping the weapons inspectors out. As Coe and Vaynman (2020) and Koblentz (2018) discuss, states may be weakened by allowing external observers to inspect their military capabilities because this information could leak and be used by adversaries. In breaking the terms of the UN resolutions, Saddam put Iraq in a better military position. Moreover, keeping inspectors out led the United States to grow concerned that Iraq was developing WMDs and was therefore a rising power. While the United States and Saddam interpreted the decision to keep weapons inspectors out in different ways, Saddam ultimately knew that his actions could provoke a military response and deliberated over how much to reveal about the state's security forces to weapons inspectors (see Woods, Palkki, and Stout 2011, 257–58; Woods et al. 2006, 15–16, 30, 91, 96–97, 125).

I do not want to gloss over the fact that Saddam was not actually building WMD and that, if the United States had all the correct information, then the United States plausibly would not have invaded Iraq. The reason the United States did not know the extent of Iraq's WMD programs was because Saddam kept weapons inspectors out. As Coe and Vaynman (2020, 349) suggest, "in principle, Iraq could be blanketed with inspectors," producing transparency in Iraq's weapons development, if not for Saddam's handling of the inspectors. Saddam's decisions created uncertainty around Iraq's WMD programs that the United States grappled with in the lead up to the 2003 invasion, setting the stage for the strategic tensions explored in research like Baliga and Sjöström (2008), Bas and Coe (2016), Debs and Monteiro (2014), Spaniel (2019), and others to operate. Ultimately, while Saddam stood to gain by turning away weapons inspectors, he did not know the United States's true willingness to accept uncertainty over Iraq's WMD programs and was too bold in his handling of weapons inspectors.

The second question to answer is whether the United States had publicly demonstrated a robust hassling capability. This was indeed the case. After the Cold War, the United States transformed much of its technical know-how in building and deploying missiles and bombers for nuclear strikes into precision strike capabilities (US Air Force 2019). Additionally, the efficacy of these capabilities and US public willingness to use them was revealed in a series of conflicts, including, but not limited to, Operation Desert Fox.

Third, there is ample evidence that Saddam's decision to refuse entry to the inspectors was influenced by US hassling capabilities. In internal discussions before the invasion, Saddam commented frequently on the US reliance on air strikes over ground operations, a sentiment echoed by his advisers (Woods et al. 2006, 15–16, 30, 96–97, 125). At one point, an

adviser was recorded in a conversation with Saddam claiming, "I believe if any incident occurs, the Americans will utilize their air strike methods, which they prefer and used recently, instead of sending troops, based on their horrific experience in Somalia" (30). Similarly, in interviews after Saddam was captured, he further stressed his belief that the United States would respond in 2003 as it had in Operation Desert Fox (Battle 2009).

Overall, Saddam knowingly chose to refuse entry to weapons inspectors to bolster his future military strength. In doing so, Saddam recognized that he risked confrontation with the United States. But, because the United States was effective at hassling, Saddam believed that the benefits of keeping weapons inspectors out outweighed the expected response of hassling and the small risk of war. However, Saddam did not know the United States's true willingness to accept his treatment of weapons inspectors, and he ultimately behaved too aggressively. While there is no way to know how Saddam would have behaved if the United States possessed a less robust hassling ability, in this counterfactual setting, it is plausible that Saddam would have taken the threat of war more seriously and would have been more open to weapons inspections.

That Saddam may have changed his behavior in response to improved US hassling capabilities (i.e., improvements that facilitate low-level conflict) in such a way that ultimately led to a war (i.e., an escalated level of conflict) has not previously been formalized or explained. While using existing theoretical frameworks—like treating improved hassling capabilities as a "failure to burn bridges" (see the appendix) or applying models like Bas and Coe (2016) and Joseph (2020)—can show that improved hassling capabilities lead to more hassling, existing research cannot explain how being better at hassling led to an escalation to war in this setting. Thus, this case presents a new perspective on the 1990s "revolution in military affairs," which viewed developments in precision strike and cyber capabilities as fundamentally altering the way conflict would be fought (Cordesman 1999). This article does not dispute that states have successfully used these technologies for hassling, plausibly in lieu of more costly and destructive wars. However, here Saddam observed the United States's increased reliance on hassling technologies and was emboldened to turn away weapons inspectors, believing that his decision would likely be met with a limited response. Ultimately, this case embodies the theoretical finding that improvements in hassling capabilities can shift the strategic decisions that revisionist states make and may ultimately result in devastating and destabilizing wars.

Predictability and the stability-instability paradox

During the Cold War, while mutually assured destruction prevented the United States and Soviet Union from engaging in nuclear war, the superpowers did compete in multiple conflict theaters at low levels, for example, by providing sympathetic rebels or governments with support and fighting proxy wars. These observations led scholars to posit a "stability-instability paradox," where stability at the nuclear level opened the possibility for competition (instability) at lower levels (Jervis 1984; Snyder 1965). While there is some empirical evidence of its existence (Early and Asal 2018; Rauchhaus 2009), the scope conditions for the paradox outside of nuclear-level stability are still an open topic (O'Neill 2019). Notably, even Snyder (1965, 189–90) hedges on the claim that nuclear stability *causes* low-level instability, emphasizing that the opposite could be true if fear of linkages or escalation between levels of conflict are too salient. This article can offer two points on these theoretical underpinnings.

First and most simply, for low-level instability to occur, states must be able and willing to engage at those low levels. Consider the United States as the defender and the USSR as the challenger. At the onset of the Cold War, funding for the newly formed Central Intelligence Agency grew rapidly, allowing the organization to participate in a wide portfolio of covert operations beyond intelligence gathering (Gaddis 2005; Snider 2015). In other words, the United States expanded its hassling capabilities, and the United States engaged in many hassling operations throughout the globe. But of course, as the model with transgression costs (see the appendix) suggests, the expanded hassling capability could have deterred transgressions. After all, if the challenger knows only well-funded and aggressive transgressions can survive a defender's robust hassling ability, then a challenger may decide transgressing is not worth the costs. Instead, the USSR was resolved enough to transgress. Essentially, the expansion in low-level conflict capabilities on both sides in the early Cold War opened the door to greater low-level instability.

Second, the predictability mechanism frames the topics of linkages and escalation within a utility analysis. The existence of low-level instability does not imply that the defender is suffering a low utility. Following improvements in hassling capabilities, the defender could be more successfully degrading transgressions at low cost while avoiding war, all of which are good outcomes. Instead, the defender experiences worse utilities when the defender becomes better at engaging in hassling, and this results in the linkages between the challenger's actions and the defender's responses becoming clearer and more predictable. When this happens, the challenger's fear of a possible escalation is undermined. If the US expansion in hassling abilities led to the United States becoming more predictable, then the USSR could more effectively "design around" the defender's deterrent threat from war and undertake aggressive, calibrated transactions without fear

of escalation (George and Smoke 1974). While a complete analysis of the Cold War is beyond the scope of this article, in some cases, there appeared to be such calibrated actions. For example, the USSR invasion of Afghanistan led to international condemnation and hassling, but there seemed little risk that the United States would commit conventional forces or engage in brinkmanship over the issue (Coll 2005). If the Soviet treatment of Afghanistan was informed by decades of US hassling operations, then it is possible the United States had become predictable.

CONCLUSION

In this article, I present a game theoretic model in which a rising power chooses how much to transgress, knowing that it faces a defending actor who can respond at multiple conflict escalation levels and who has private information about its willingness to engage at different levels. How do the defender's capabilities affect its outcomes? When the defender improves its ability to engage in low-level conflict, two mechanisms can cause worse outcomes for the defender: the rising power may be emboldened, or the defender may become predictable. In short, these results arise when improved low-level conflict capabilities undermine the defender's ability to effectively use its private information.

The results here suggest that political scientists and policy makers need to take a harder look at low-level conflict capabilities. The logic of having "many tools in the policy toolbox" can be counterproductive because of the strategic responses these tools can produce in rivals. While this article does not suggest that having many policy options is always bad, it does suggest that having more tools can, under some conditions, lead to systematically worse outcomes. More research is needed on this topic, especially work that considers uncertainty over willingness to engage in multiple kinds of conflict. Future work should also take a harder look at the relationship between war and hassling capabilities and outcomes, specifically understanding the scope of when better capabilities can lead to worse outcomes. Finally, while there is outstanding work on the relationship between conflict capabilities and likelihood of war, future work should further examine the relationship between (all kinds of) conflict capabilities and utility outcomes both within the standard deterrence framework and within models where actors can degrade the other side at any time.

ACKNOWLEDGMENTS

Thanks to Brad Smith, Andrew Coe, Brenton Kenkel, Kris Ramsay, William Spaniel, James Fearon, Andrew Little, Erik Gartzke, Michael Joseph, Braden Sopher, David Siegel, Cassy Dorff, Heesun Yoo, and seminar participants at the FMIR, Southern Political Science Association, Duke, the Virtual Formal Theory Workshop, and Vanderbilt. I also thank three anonymous reviewers for their feedback. I am indebted to Haonan Dong, Ethan Sun, and the Vanderbilt ROCCA lab for outstanding research assistance. All errors belong to the author.

REFERENCES

Acemoglu, Daron, and James A. Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York: Crown.

Acharya, Avidit, and Edoardo Grillo. 2015. "War with Crazy Types." *Political Science Research and Methods* 3 (2): 281–307.

Ashworth, Scott, and Ethan Bueno de Mesquita. 2006. "Monotone Comparative Statics for Models of Politics." *American Journal of Political Science* 50:214–31.

Baliga, Sandeep, Ethan Bueno de Mesquita, and Alexander Wolitzky. 2020. "Deterrence with Imperfect Attribution." *American Political Science Review* 114:1155–78.

Baliga, Sandeep, and Tomas Sjöström. 2008. "Strategic Ambiguity and Arms Proliferation." *Journal of Political Economy* 116:1023–57.

Baliga, Sandeep, and Tomas Sjöström. 2020. "The Strategy and Technology of Conflict." *Journal of Political Economy* 128:3186–219.

Bas, Muhammet A., and Andrew J. Coe. 2012. "Arms Diffusion and War." *Journal of Conflict Resolution* 56:651–74.

Bas, Muhammet A., and Andrew J. Coe. 2016. "A Dynamic Theory of Nuclear Proliferation and Preventive War." *International Organization* 70:655–85.

Battle, Joyce. 2009. "Saddam Hussein Talks to the FBI." National Security Archive.

Benson, Brett V., Adam Meirowitz, and Kristopher W. Ramsay. 2016. "Changing Capabilities, Uncertainty, and the Risk of War in Crisis Bargaining." *Research and Politics* 3.

Benson, Brett V., and Bradley C. Smith. 2020. "Do Alliances Cause Commitment Problems and War?" Working paper.

Blainey, Geoffrey. 1988. *Causes of War*. New York: Simon & Schuster.

Buchanan, Ben. 2020. *The Hacker and the State: Cyber Attacks and the New Normal of Geopolitics*. Cambridge, MA: Harvard University Press.

Chassang, Sylvain, and Gerard Padró i Miquel. 2009. "Defensive Weapons and Defensive Alliances." *American Economic Review* 99:282–86.

Chassang, Sylvain, and Gerard Padró i Miquel. 2010. "Conflict and Deterrence under Strategic Risk." *Quarterly Journal of Economics* 125:1821–58.

Coe, Andrew J. 2018. "Containing Rogues: A Theory of Asymmetric Arming." *Journal of Politics* 80 (4): 1197–210.

Coe, Andrew J., and Jane Vaynman. 2020. "Why Arms Control Is So Rare." *American Political Science Review* 114 (2): 342–55.

Coll, Steve. 2005. *Ghost Wars: The Secret History of the CIA, Afghanistan, and Bin Laden, from the Soviet Invasion to September 10, 2001*. New York: Penguin.

Copeland, Dale C. 2001. *The Origins of Major War*. Ithaca, NY: Cornell University Press.

Cordesman, Anthony H. 1999. "The Lessons of Desert Fox: A Preliminary Analysis." Center for Strategic and International Studies, February 16.

Cruz, José Miguel, and Angélica Durán-Martínez. 2016. "Hiding Violence to Deal with the State: Criminal Pacts in El Salvador and Medellin." *Journal of Peace Research* 53 (2): 197–210.

Debs, Alexandre, and Nuno P. Monteiro. 2014. "Known Unknowns: Power Shifts, Uncertainty, and War." *International Organization* 68:1–31.

Di Lonardo, Livio, and Scott A. Tyson. 2022. "Political Instability and the Failure of Deterrence." *Journal of Politics* 84 (1): 180–93.

Duelfer, Charles. 2004. *Comprehensive Report of the Special Advisor to the DCI on Iraq's WMD*, vol. 1. Washington, DC: Central Intelligence Agency.

Early, Bryan R., and Victor Asal. 2018. "Nuclear Weapons, Existential Threats, and the Stability-Instability Paradox." *Nonproliferation Review* 25:223–47.

Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88:577–92.

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

Fearon, James D. 1997. "Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs." *Journal of Conflict Resolution* 41:68–90.

Fey, Mark, and Kristopher W. Ramsay. 2011. "Uncertainty and Incentives in Crisis Bargaining: Game-Free Analysis of International Conflict." *American Journal of Political Science* 55:149–69.

Fuhrmann, Matthew, and Sarah E. Kreps. 2010. "Targeting Nuclear Programs in War and Peace: A Quantitative Empirical Analysis, 1941–2000." *Journal of Conflict Resolution* 54:831–59.

Gaddis, John Lewis. 2005. *Strategies of Containment*. Oxford: Oxford University Press.

Gannon, Andres J., Erik Gartzke, Jon R. Lindsay, and Peter Schram. 2020. "The Shadow of Deterrence." Unpublished manuscript.

Gartzke, Erik, and Jon R. Lindsay. 2017. "Thermonuclear Cyberwar." *Journal of Cybersecurity* 3:37–48.

George, Alexander L., and Richard Smoke. 1974. *Deterrence in American Foreign Policy: Theory and Practice*. New York: Columbia University Press.

Gurantz, Ron, and Alexander V. Hirsch. 2017. "Fear, Appeasement, and the Effectiveness of Deterrence." *Journal of Politics* 79:1041–56.

Horowitz, Michael, and Dan Reiter. 2001. "When Does Aerial Bombing Work? Quantitative Empirical Tests, 1917–1999." *Journal of Conflict Resolution* 45:147–73.

Huth, Paul K. 1999. "Deterrence and International Conflict: Empirical Findings and Theoretical Debates." *Annual Review of Political Science* 2:25–48.

Jervis, Robert. 1984. *The Illogic of American Nuclear Strategy*. Ithaca, NY: Cornell University Press.

Joseph, Michael. 2020. "The Secret Failure of Nuclear Containment." Unpublished manuscript.

Katz, Yaakov. 2010. "And They Struck Them with Blindness." *Jerusalem Post*.

Koblentz, Gregory D. 2018. "Saddam versus the Inspectors: The Impact of Regime Security on the Verification of Iraq's WMD Disarmament." *Journal of Strategic Studies* 41:372–409.

Kreps, Sarah E., and Matthew Fuhrmann. 2011. "Attacking the Atom: Does Bombing Nuclear Facilities Affect Proliferation?" *Journal of Strategic Studies* 34:161–87.

Kydd, Andrew. 2000. "Arms Races and Arms Control: Modeling the Hawk Perspective." *American Journal of Political Science* 44 (2): 228–44.

Kydd, Andrew H., and Roseanne W. McManus. 2017. "Threats and Assurances in Crisis Bargaining." *Journal of Conflict Resolution* 61:325–48.

Lanoszka, Alexander. 2016. "Russian Hybrid Warfare and Extended Deterrence in Eastern Europe." *International Affairs* 92:175–95.

Levy, Jack S. 2011. "Preventive war: Concept and propositions." *International Interactions* 37:87–96.

Lindsay, Jon R., and Erik Gartzke. 2018. "Coercion through Cyberspace: The Stability-Instability Paradox Revisited." In Kelly M. Greenhill and Peter Krause, eds., *Coercion: The Power to Hurt in International Politics*. New York: Oxford University Press.

Mazarr, M. J. 2015. "Struggle in the Gray Zone." War on the Rocks, December 22.

McCormack, Daniel, and Henry Pascoe. 2017. "Sanctions and Preventive War." *Journal of Conflict Resolution* 61:1711–39.

McMahon, R. Blake, and Branislav L. Slantchev. 2015. "The Guardianship Dilemma: Regime Security through and from the Armed Forces." *American Political Science Review* 109:297–313.

Mehta, Rupal N., and Rachel Elizabeth Whitlark. 2017. "The Benefits and Burdens of Nuclear Latency." *International Studies Quarterly* 61:517–28.

Meirowitz, Adam, Massimo Morelli, Kristopher W. Ramsay, and Francesco Squintani. 2019. "Dispute Resolution Institutions and Strategic Militarization." *Journal of Political Economy* 127:378–418.

Meirowitz, Adam, and Anne E. Sartori. 2008. "Strategic Uncertainty as a Cause of War." *Quarterly Journal of Political Science* 3:327–52.

Morrow, James D. 1989. "Capabilities, Uncertainty, and Resolve: A Limited Information Model of Crisis Bargaining." *American Journal of Political Science* 33 (4): 941–72.

Myerson, Roger B. 1979. "Incentive Compatibility and the Bargaining Problem." *Econometrica* 47 (1): 61–73.

O'Neill, Barry. 2019. "A Game Theory Analysis of the Stability-Instability Paradox." In Lauren Ice, James Scouras, Kelly Rooker, Robert Leonhard, and David McGarvey, eds., *Game Theory and Nuclear Stability in Northeast Asia*. Baltimore: Johns Hopkins University Applied Physics Laboratory, 68–73.

Paine, Jack. 2021. "The Dictator's Power-Sharing Dilemma: Countering Dual Outsider Threats." *American Journal of Political Science* 65 (2): 510–27.

Post, Abigail. 2019. "Flying to Fail: Costly Signals and Air Power in Crisis Bargaining." *Journal of Conflict Resolution* 63:869–95.

Powell, Robert. 1989. "Nuclear Deterrence and the Strategy of Limited Retaliation." *American Political Science Review* 83 (2): 503–19.

Powell, Robert. 1999. *In the Shadow of Power: States and Strategies in International Politics*. Princeton, NJ: Princeton University Press.

Powell, Robert. 2006. "War as a Commitment Problem." *International Organization* 60:169–203.

Powell, Robert. 2015. "Nuclear Brinkmanship, Limited War, and Military Power." *International Organization* 69:589–626.

Ramsay, Kristopher W. 2017. "Information, Uncertainty, and War." *Annual Review of Political Science* 20:505–27.

Rauchhaus, Robert. 2009. "Evaluating the Nuclear Peace Hypothesis: A Quantitative Approach." *Journal of Conflict Resolution* 53:258–77.

Reiter, Dan. 2005. "Preventive Attacks against Nuclear Programs and the 'Success' at Osiraq." *Nonproliferation Review* 12:355–71.

Ritter, Emily Hencken. 2014. "Policy Disputes, Political Survival, and the Onset and Severity of State Repression." *Journal of Conflict Resolution* 58:143–68.

Salanié, Bernard. 2005. *The Economics of Contracts: A Primer*. Cambridge, MA: MIT Press.

Schelling, Thomas C. 1966. *Arms and Influence*. New Haven, CT: Yale University Press.

Schelling, Thomas C. 1980. *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

Schram, Peter. 2019. "Managing Insurgency." *Journal of Conflict Resolution* 63:2319–53.

Schram, Peter. 2021. "Hassling: How States Prevent a Preventive War." *American Journal of Political Science* 65:294–308.

Schultz, Kenneth A. 2010. "The Enforcement Problem in Coercive Bargaining: Interstate Conflict over Rebel Support in Civil Wars." *International Organization* 64:281–312.

Shapiro, Jacob N. 2013. *The Terrorists Dilemma: Managing Violent Covert Organizations*. Princeton, NJ: Princeton University Press.

Slantchev, Branislav L., and Ahmer Tarar. 2011. "Mutual Optimism as a Rationalist Explanation of War." *American Journal of Political Science* 55:135–48.

Smith, Alastair. 1998. "International Crises and Domestic Politics." *American Political Science Review* 92:623–38.

Snider, L. Britt. 2015. *The Agency and the Hill*. Washington, DC: GPO.

Snyder, Glenn H. 1965. "The Balance of Power and the Balance of Terror." In Frederick H. Hartmann, ed., *World in Crisis*. New York: Macmillan, 180–91.

Spaniel, William. 2019. *Bargaining over the Bomb: The Successes and Failures of Nuclear Negotiations*. Cambridge: Cambridge University Press.

Spaniel, William. 2020. "Power Transfers, Military Uncertainty, and War." *Journal of Theoretical Politics* 32:538–56.

Spaniel, William, and Iris Malone. 2019. "The Uncertainty Trade-Off: Reexamining Opportunity Costs and War." *International Studies Quarterly* 63:1025–34.

Trager, Robert F. 2010. "Diplomatic Calculus in Anarchy: How Communication Matters." *American Political Science Review* 104 (2): 347–68.

Trejo, Guillermo, and Sandra Ley. 2021. "High-Profile Criminal Violence: Why Drug Cartels Murder Government Officials and Party Candidates in Mexico." *British Journal of Political Science* 51:203–29.

US Air Force. 2019. "AGM-86B/C/D Missiles." Technical report.

Wagner, R. Harrison. 2000. "Bargaining and War." *American Journal of Political Science* 44 (3): 469–84.

Waltz, Kenneth N. 1979. *Theory of International Politics*. New York : McGraw-Hill.

Woods, Kevin M., David D. Palkki, and Mark E. Stout. 2011. *The Saddam Tapes: The Inner Workings of a Tyrant's Regime, 1978–2001*. Cambridge: Cambridge University Press.

Woods, Kevin M., Michael R. Pease, Mark E. Stout, Williamson Murray, and James G. Lacey. 2006. "Iraqi Perspectives Project: A View of Operation Iraqi Freedom from Saddam's Senior Leadership." Technical report, US Joint Forces Command, Norfolk, VA.