When Deterrence Fails: How Improved Hassling Capabilities Produce Worse Outcomes*

Peter Schram

June 24, 2020

Abstract

I formalize interactions between an endogenously rising state and a rival, defending state that can accept the rising state's rise, go to war before the rise comes to fruition, or degrade the rising state's growth through low-level conflict operations that I call "hassling." The novelty here is that the defending state has private information about its hassling capabilities; this implies that the rising state does not know how much it can rise without provoking the defending state to hassle or go to war. I find that when the defending state's ability to hassle improves, it can provoke strategic responses in the rising state that undermines the defending state's ability to use its private information productively and produces lower utility for the defending state. This model provides insight into Saddam Hussein's decision-making leading up to the 2003 U.S. invasion and the proliferation of proxy conflicts during the Cold War.

*Thanks to Brad Smith, Andrew Coe, Brenton Kenkel, Kris Ramsay, William Spaniel, Michael Joseph, Braden Sopher, CPASS, and the Vanderbilt Conflict Workshop and Faculty workshop for helpful comments. I am indebted to Haonan Dong, Ethan Sun, and the Vanderbilt ROCCA lab for outstanding research assistance.

Leading up to 2003, Saddam Hussein repeatedly denied entry to United Nations 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 did have good reason to keep inspectors out: he believed that the weapons inspectors would document details on the extent and location of his armaments, information that could be used maliciously by Iraq's adversaries (Coe and Vaynman, 2020). Still, the 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 U.S. forces. And, transcripts from the Saddam regime suggested that Iraq knew that its behavior around WMD issues put it at risk for a confrontation with the United States. Indeed, the well-known logic of commitment problems identifies this scenario as one in which preventive war is likely (Fearon, 1995; Powell, 2006; Debs and Monteiro, 2014)—which is ultimately what happened. The puzzle, then, is why was Saddam Hussein not deterred by the might of United States 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 was willing to gamble on keeping weapons inspectors out because he did not expect he would face the full might of the United States military. In 1998, in response to Iraq turning away weapons inspectors, the U.S. conducted Operation Desert Fox, a four-day bombing campaign against Iraqi weapons facilities. In this operation, the U.S. illustrated that it had the ability and willingness to use targeted low-level conflict operations in response to rising powers. If Saddam believed the U.S. 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 "hassling" operations such as limited bombing campaigns, the U.S. may have undermined its own 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, at the time, a success—they may provoke problematic strategic 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.

In this paper, I formalize the above intuition. I consider a rising state that chooses how much to “invest in a rising technology" that will make the rising state more powerful in the future (for example, by investing in a nuclear weapons program). In response, a rival defending state can accept the investment, initiate a preventive war to stop the investment from coming to fruition, or "hassle," which, following Schram (Forthcomming), is the use of limited conflict to degrade a rising state's investment in rising technology while allowing bargaining to continue. What distinguishes this work from substantively similar research (Bas and Coe, 2016; Spaniel, 2019; Joseph, 2020) is that the defending state's hassling abilities have both a public and private component. The rising state uses its observations of the defending state's public hassling capabilities in making its investment decision, but it does not observe the defending state's exact private willingness to hassle. Thus, the rising state does not know how much it can invest in the rising technology before it provokes hassling or war. In this setting, improved hassling capabilities can in some situations produce worse outcomes—more war, more hassling, and higher levels of rising technology—which result in lower utility for the defending state. All else equal, being better at hassling can encourage a rising state to take undesirable actions in the eyes of the defending state, thus producing a deterrence failure.

Improvements in hassling capabilities produce a deterrence failure in one of two ways. First, improvements in the defending state's hassling capabilities can produce a deterrence failure by *embolden the rising state*. This happens because the rising state now believes it can undertake valuable, potentially escalatory actions without running as high a risk of a devastating war. When Iraq chose to turn away weapons inspectors and lie about its capabilities, it did so knowing that the United States may respond with military force. And, while war was a possible military response, because the U.S. was good at hassling (as evidenced by, for example, Operation Desert Fox), Iraq expected hassling as the most likely response, and therefore determined that the benefits of turning away weapons inspectors out outweighed the risks of the expected military response. In a counterfactual setting where the U.S. was worse at hassling, Iraq might have expected that refusing entry to weapons 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 to inspectors. In other words, because the U.S. was good at hassling, it undermined war as a deterrent threat, and emboldened Iraq to gamble on keeping weapons inspectors out.

Second, improvements in the defending state's hassling capabilities can produce a deterrence failure by *making the defending state more predictable.* In the model here, private information is useful because it allows the player with private information to posture to attain better outcomes. The threat of war is an effective deterrent; if a rising state thinks that building a nuclear bomb will provoke its rivals to declare war today, then the rising state will not invest in the reactor. When the defending state postures effectively, it convinces the rising state that the defending state would go to war over an investment in rising technology when, in fact, the defending state would not have gone to war. Under some conditions, improvements in the defending state's hassling capabilities diminish its ability to posture effectively, thus allowing the rising state to calibrate its investments in rising technology to extract more bargaining surplus from the defending state. I examine this mechanism in the context of the proxy conflicts that occurred during the Cold War and the stability-instability paradox.

This is not the first paper to suggest that an improved capability to conduct low-level conflict or precision strikes against nuclear facilities can be counter-productive for the state making these improvements (Schelling, 1980; Powell, 2015; Bas and Coe, 2016; Joseph, 2020). However, this paper makes three important contributions. First, this model formalizes an important and previously unexplored strategic dynamic in which a rising state wants to invest in a rising technology but does not know how much it can invest before its opponents will respond with hassling or war. This dynamic is distinct from that explored in substantively similar work like Debs and Monteiro (2014), Bas and Coe (2016), and Spaniel (2019, pp. 244-289).[1] Second, this is the first paper to identify cases of deterrence failure stemming from the defending state becoming more predictable and the rising state becoming emboldened. Third, I show, for a broad class of models, that the predictability and emboldening mechanisms are not just two ways deterrence failures can occur, but rather they are the *only* ways deterrence failures can occur.

Hassling is a pervasive feature of the international system. States have used hassling to degrade power shifts stemming from investment in nuclear weapons (Operation Outside the Box, and the Stuxnet computer worm), future alliances (the Quemoy Crisis and Russian involvement in Ukraine), and annexing valuable territory (U.S. support for Afghan mujahideen to fight the Soviet Union).[2] And, previous work has demonstrated that hassling can be useful in international crises by offering a less costly alternative to preventive wars (Coe, 2018).[3] This paper confirms that hassling can be useful in a particular crisis, but it also shows that building a strong hassling capability can invoke aggressive strategic responses in adversaries. And, as I describe below, this formalization offers insights into the proxy wars fought during the Cold War and into Saddam's decision making in the lead up to the 2003 U.S. invasion.

[1] I discuss this in more detail below, but as one difference, this is the only paper to consider low-level conflict as a distinct policy option from war.

[2] These examples (and others) are discussed in Schram (Forthcomming) and Benson and Smith (2020).

[3] As McCormack and Pascoe (2017) show, similar results can apply to sanctions acting as alternatives to preventive wars.

I proceed as follows. In Section 1 I outline the theory. In Section 2 I present the model, and in Section 3 I provide the equilibrium outcomes. Section 4 describes when improvements in hassling capabilities produce a deterrence failure. Section 5 illustrates the mechanisms at work in two cases. I describe additional results and extensions in Section 6, and I conclude in Section 7.

1 Theoretical Background

1.1 Key Concepts and Terminology

This paper will reference several types of technologies and political behaviors. In this section, I provide definitions of these concepts and illustrate them using Operation Outside the Box. Operation Outside the Box was a hassling operation that Israel undertook in 2007 in response to the discovery that Syria was building a nuclear reactor in Al Kibar. In response, Israel used a cyberattack to disable Syrian air defenses and conducted an airstrike on the reactor (Harel and Benn, 2018). Because building a reactor moved Syria closer to being able to produce a nuclear weapon, this was a clear investment in a "rising technology."

Definition: *States invest in* ***rising technologies*** *by pursuing opportunities that strengthen their future military capabilities.*

If investments in rising technologies are allowed to come to fruition, the state making the investment will have improved wartime abilities in future conflicts and more leverage in future negotiations. Under this definition, there is nothing inherently scientific about rising technologies. While the concept of investing in rising technologies can apply to investing in military nuclear, space, or cyber technologies (Gartzke and Lindsay, 2017), it can also apply to a range of political behaviors including amassing conventional forces (Copeland, 2001), forming alliances (Benson and Smith, 2020), and securing geopolitically valuable territory (Fearon, 1996; Powell, 2006).

By investing in a rising technology, the rising state creates a commitment problem. In response, rival states may initiate a preventive war so that the rising technology does not come to fruition (Fearon, 1995; Powell, 2006; Bas and Coe, 2012; Debs and Monteiro, 2014; Bas and Coe, 2016). But rival states have another option short of war: they can "hassle" the rising state. When Israel conducted a cyberattack and airstrike against Syria to degrade Syria's nuclear aspirations, it engaged in "hassling."

Definition: Hassling *is the limited use of costly and destructive military assets against another state with the intent of blunting power shifts to allow for bargaining to occur.*

Hassling was previously defined and discussed in Schram (Forthcomming). Hassling operates like a steam valve: in an international system where a rising power may provoke other states to declare a preventive war, hassling can diffuse the situation at a lower cost to both parties than war. While war halts or destroys a rising technology and resolves the political crisis, hassling degrades the rising technology while still allowing for negotiations to occur.[4] The 2007 strike against the Syrian reactor at Al Kibar, an example of hassling, was a blow to the Syrian nuclear program, but neither resolved the host of political disputes between Syria and Israel nor prevented Syria from future investments in a nuclear program. Operation Outside the Box is therefore different from the 2003 invasion of Iraq, a preventive war. The 2003 invasion was conducted to insure that Iraq could not develop WMD, thus destroying—not blunting—Iraq's future power shift, and the invasion eliminated the possibility of future peaceful negotiations with the Baathist government by overthrowing it. Many instances of

[4] In the standard "costly lottery" treatment of war, war occurs then ends with one side acquiring the political asset (Fearon, 1995).

limited strikes against nuclear facilities (Reiter, 2005; Fuhrmann and Kreps, 2010),[5] hybrid conflict (Lanoszka, 2016), and gray-zone conflict (Mazarr, 2015; Votel *et al.*, 2016) could qualify as hassling. Hassling operates similarly to how other works treat arming, sanctions, and containment regimes (McCormack and Pascoe, 2017; Coe, 2018; Joseph, 2020).

For Israel to conduct Operation Outside the Box, it relied on its "hassling capabilities."

Definition: Hassling capabilities *are the tools that a state uses and willingness of a state to hassle.*

My project in this paper is to examine the equilibrium effects of improvements in hassling capabilities. Bas and Coe (2016) and Joseph (2020) also both consider this substantive topic. The key distinction here is that there is a private component to the defending state's hassling capabilities and, thus, its willingness to hassle. As a result, the rising state is uncertain over how much it can invest in rising technologies 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 states know for certain about the defending state's hassling capabilities, which largely comes from military intelligence and observations of a defending state's previous hassling operations. For example, in 2007, Syria knew that Israel had a robust hassling capability, as Syria had observed the 1981 Operation Opera and more recent Israeli limited operations in Lebanon and Palestinian territories. The private component can be interpreted as what other states do not know for certain about the defending state's hassling capabilities, which largely comes from uncertainty over the defending state's willingness to hassle or the defending state's secret technological capabilities.

[5] As discussed in Schram (Forthcomming), some contemporary definitions of preventive war are broad enough to include some instances of hassling (Levy, 2011). However, there is value in treating hassling as a distinct policy choice to preventive war, as there are some technologies that lend themselves better to limited, low-level operations than to more costly, conventional forms of escalation.

To continue the example, Syria likely did not know that Israel possessed a cyberweapon that could disrupt its air defenses. Furthermore, even if Syria had some idea that Israel possessed such a cyberweapon, Syria did not know Israel's willingness to use the weapon in a strike on the Al Kibar reactor.[6] Of course, the idea that a state possesses private information about its ability to engage in conflict is not new, and the claim that states have private information about their hassling capabilities is analogous to claims that states have private information about their ability to wage war (Fearon, 1995; Fey and Ramsay, 2011; Spaniel, 2019).

Following this logic, when I refer to public improvements in hassling capabilities, I mean improvements to the defending state's ability to hassle that can be observed by rivals. These improvements could occur through publicly announced upgrades in weapons capabilities, military training exercises for hassling-type operations, or even strategic leaking of classified data on capabilities.

In my model, I consider two types of defending states: high hassling capacity states and low hassling capacity states. I assume states are not able to tailor their hassling capabilities to the crisis at hand. This assumption has empirical justification as it is difficult to vary hassling capabilities in the lead up to a political crisis. For example, during Operation Outside the Box, the Israeli Air Force deployed F-15Is and F-16Is, systems that Israel acquired in the 1990s, a decade earlier. A state's hassling capabilities at the time of a specific political crisis are subject to a wide range of factors, including past conflicts, delays in weapons development or procurement, and domestic politics. That said, even if this feature is viewed as incorrect, this paper still produces a surprising result: states may rationally turn down costless opportunities to improve their hassling capabilities due to the problematic strategic response it can induce in adversaries.

[6] Because once a cyberattack is detected the exploit can be fixed, defending states may be unwilling to use an exploit in one case because then it could not be used it in another case (Gartzke and Lindsay, 2015).

Operation Outside the Box was a success. But the fact that Israel had to conduct Operation Outside the Box represents a kind of deterrence failure. For Israel, ideally, Syria would have never invested in the reactor in the first place, and Israel would not have needed to conduct a costly and risky hassling campaign.

Definition: *A state experiences a* ***deterrence failure*** *when a state fails to prevent an opponent from undertaking undesirable activities, resulting in the state experiencing overall worse outcomes.*

Following classical deterrence theory, when a state prevents an opponent from undertaking an action through the threat of retaliation, the state has deterred its opponent. When an opponent is undeterred from taking an action that the state dislikes, the state experiences a deterrence failure. It is worthwhile highlighting that this definition does not identify a deterrence failure in terms of conflict. For example, assuming Israel does not wish Syria to become a nuclear state, then Israel experiences a deterrence failure if Syria builds a nuclear reactor and then a nuclear bomb and Israel does nothing to stop this from occurring. I formalize this below.

1.2 Theory

I consider the following interaction. One state, a rising power, chooses how much to invest in its rising technologies. If the investments are allowed to come to fruition, the rising state can leverage its increased capabilities into better future political outcomes. However, upon observing the investment, a rival defending state has three choices: it can accept the rising state's investment; it can go to war to decisively challenge the rising state before the capabilities come to fruition; or it can hassle to degrade the investments. Because this response depends on the defending state's hassling capabilities (which has a private component), the rising state does not know how the defending state will respond to a selected level of investment. I formalize this interaction and consider what effect improvements in hassling capabilities have on the defending state. I find that better hassling capabilities can hurt the defending state through two mechanisms: improvements in public hassling capabilities may make the defending state more predictable or may embolden the rising state.

How can improvements in hassling technologies make the defending state more predictable? When private hassling capabilities (as opposed to public hassling capabilities) play a large role in determining the defending state's overall hassling capabilities, the defending state's response to a selected investment in rising technology is difficult to predict. This unpredictability can be valuable to the defending state; because it is difficult to know how the defending state will behave, the rising state may scale back its investment in the rising technology to avoid war. This is akin to the defending state posturing or bluffing to get the rising state to reduce its investments. However, when the defending state publicly becomes better at hassling, sometimes the rising state has more information about how the defending state will respond; following the defending state becoming better at hassling, the rising state can at times better predict when their investments would be met with hassling and not war. For example, because the United State had an robust ability to conduct low-level covert operations and because a war would be costly, the Soviet Union knew that their invasion of Afghanistan in 1979 would most likely not be met with war, but could be met with hassling. When public improvements in hassling capabilities reduce the uncertainty about how the defending state will respond, the defending state cannot effectively posture, undermining the deterrent threat from war. Essentially, improvements in public hassling capabilities can diminish the non-rising state's benefits from its private information.[7]

How can improvements in hassling technologies embolden a rising state? When choosing the amount to invest in rising technology, the rising state faces a trade-off: a larger invest-

[7]In the Online Appendix, I offer an alternate perspective on the predictability mechanism through the lens of poker.

ment is advantageous if the defending state does not declare war, but a larger investment increases the likelihood that the defending state will declare war. The terms of this trade-off—whether a small increase in investment will produce a small or large increase in the likelihood of war—is dictated by the defending state's hassling capabilities, that has both public and private components. Under some trade-off terms—say an 8% increase in rising technology investment produces a 20% increase in the likelihood of war—the rising state is unwilling to increase its investment in rising technology. However, under different trade-off terms that could occur when the defending state is better at hassling—a 27% increase in rising technology produces a 20% increase in the likelihood of war—the rising state would be willing to increase its investment.[8] If the rising state is emboldened in this way, being better at hassling can actually cause worse outcomes for the defending state to an extent that offsets its gains from being better at hassling.

What are the empirical implications of these results? In short, under select conditions, if the defending state is better at hassling, rivals will act more aggressively in their pursuit of a rising technology, resulting in a greater likelihood of war, more pervasive and intensive hassling campaigns, or some combination of both. For example, in the aftermath of the Cold War, the United States possessed a robust hassling capability that was further bolstered by transitioning several nuclear weapon delivery systems to use in conventional settings.[9] If the conditions for emboldening or predictability held, as a result of these improvements in hassling capabilities, a number of revisionist states may have been emboldened to pursue nuclear technologies or engage in other revisionist moves, compelling the United States and its allies to engage in hassling and even war to prevent proliferation. It can be imagined that in a counterfactual universe in which the United States invested less in hassling and more in conventional military strategy, perhaps those states would have been more deterred

[8]These numbers follow from the parameters used in Figure 2.

[9]Many of the weapons systems that used in Operation Desert Fox, like the AFM-86C cruise missile, B-1 Bomber, and B-52 Stratofortress, were developed during the Cold War and later retrofit to deploy non-nuclear payloads (U.S. Air Force, 2019a,c,b).

by the threat of war and would not have pursued revisionist policy to the extent that they did.

Importantly, this paper also shows that becoming better at hassling does not always produce a deterrence failure. For example, it could be that a defending state improves its hassling capabilities by developing a wide range of hassling technologies that could be used separately or together, with a range of possible complementarities between technologies. If this is the case, then it is possible that the predictability conditions will not hold, and improvements in hassling capabilities can lead to better outcomes for the defending state. Alternatively, if a defending state becomes so effective at hassling that it could degrade a rival's investment at almost no cost, then the rising state would be deterred from investing in their rising technology. Of course, achieving this degree of hassling efficacy may be difficult, as rival countries may be willing to absorb huge costs to develop certain technologies, such as nuclear weapons[10] or could discover countermeasures to shield themselves from hassling. Another way that improvements in hassling capabilities could improve outcomes for the defending state is if those improvements also improve its wartime capabilities, which is discussed in Section 6.1.

1.3 Related Theory

While other works have illustrated the result that conflict can arise following improvements in low-level conflict capabilities (Schelling, 1980; Powell, 2015; Bas and Coe, 2016), this paper makes two theoretical contributions to our understanding of this result. First, this is the only paper to examine interactions where a rising power that can choose how fast it rises is uncertain over its opponent's willingness to conduct low-level conflict, and where war is also an option. Because this setting is critical for the mechanisms to function, the results here are theoretically novel. Second, this paper not only formalizes the predictability and emboldening mechanisms and shows when they arise, but it also shows, for a broad

[10] For example, in 1965, Pakistan's Prime Minister Zulfiqar Ali Bhutto said Pakistanis "will eat grass and leaves" to pay for a nuclear bomb (Anderson and Khan, 1998).

class of models, that these are the *only* two mechanisms that can produce a deterrence failure.

A now vast literature analyzes power shifts and preventive wars (Levy, 1987; Fearon, 1995; Powell, 2006). A subset of these considers preventive wars with an endogenously rising power (Debs and Monteiro, 2014; Bas and Coe, 2016; Spaniel, 2019; Meirowitz *et al.*, 2019) or with a state deliberately making some revisionist action (Schultz, 2010). These models share a common feature: the defending state cannot perfectly observe what investments the rising state is taking and so must base its strategy on an imperfect signal. In contrast, I treat investments in rising technology as commonly observed. While some investments in rising technology may be hidden, many, like amassing conventional forces or securing valuable territory, are quite observable. And, even when one state is investing in nuclear technology, which understandably is often the first "rising technology" that comes to mind, the defending state is often aware of the investment. In Operation Opera, Operation Outside the Box, and the deployment of the Stuxnet Worm, enough was known about the respective nuclear programs to inform the decision to hassle and to conduct precise, targeted operations. Finally, even when investments in rising technology send stochastic signals to rivals, the central tension explored here—the state investing in a rising technology is uncertain of how its opponent will respond to the true investment level—continues to exist.

This uncertainty on the part of the rising state about how its opponent will respond to its investment contributes to the broader literature suggesting that states face uncertainty over their opponent's military capabilities (Morrow, 1989; Fearon, 1995; Fey and Ramsay, 2011; Slantchev and Tarar, 2011; Spaniel and Bils, 2018).[11] This paper models a specific type of uncertainty, namely uncertainty over a state's ability and willingness to use hassling, which is new. This feature also makes this work distinct from related research that considers how changes in wartime capabilities affects outcomes (like utilities and likelihood of war)

[11] See Ramsay (2017) for an excellent review.

(Benson *et al.*, 2016).

In all the models cited above, when a state is dissatisfied with its outcome, it declares war. Of course, the real world is not this simple and there are many possible destructive policy responses that decision makers select from, as highlighted in non-formal works like Levy (2011), Gartzke and Lindsay (2017), and Mehta and Whitlark (2017). An emerging branch of formal research considers a policymaker who faces a range of policy outcomes, including limited war (Powell, 2015) and efforts that degrade a rising power's rise (McCormack and Pascoe, 2017; Coe, 2018; Joseph, 2020). This literature often finds that low-level conflict can be a useful alternative to preventive war. However, these works do not consider an endogenously rising state that is uncertain over how fast it can rise. This is the first paper to show that the defending state's high capacity to hassle can interact with its private information in ways that encourage the rising state to invest more aggressively in their rising technology.

2 Model

2.1 Characterizing Hassling Games

I consider two states, A and D, that are in a "hassling game" over a divisible asset with a normalized value of 1. In the hassling game, A wants to invest in a rising technology (denoted $t \in \mathcal{T} \subseteq \mathbb{R}_{\geq 0}$)[12] to improve its future wartime capabilities. In response to the selected t,[13] D will accept the rising technology, engage in hassling to degrade the rising technology (denoted $h \in \mathcal{H} \subseteq \mathbb{R}_{\geq 0}$), or go to war to prevent the rising technology from coming to fruition. When D does not declare war, A's investment t and D's hassling level h affects A's future likelihood of winning in war, which affects future bargaining between A and D.

[12] I define $\mathbb{R}_{\geq 0} = \{a : a \geq 0\}$.

[13] I treat t as commonly observed, an assumption that I describe in Section 1.3.

Critically, D's costs from hassling consist of a public parameter $\alpha \in \mathcal{A} \subseteq \mathbb{R}_{\geq 0}$ and a private type $\theta \in \Theta \subseteq \mathbb{R}_{\geq 0}$ known only to D. The rationale for this feature is described in detail in Section 1.1, but to offer one interpretation, α can be thought of as D's observed hassling capacity (i.e., what D has done in the past), and θ could be thought of as D's willingness to engage in hassling. In this setup, A does not know how much it can invest before D will start hassling or go to war.

I outline the game form below. In the subsequent analysis, I will compare outcomes to the game across low and high public hassling capabilities, or that is, across set elements $\{\underline{\alpha}, \bar{\alpha}\} \subseteq \mathcal{A}$, with $\underline{\alpha} < \bar{\alpha}$.

1. I let $\{\underline{\theta}, \bar{\theta}\} = \Theta$ with $\underline{\theta} < \bar{\theta}$. Nature sets $\theta = \underline{\theta}$ with probability $Pr(\underline{\theta})$ and sets $\theta = \bar{\theta}$ with probability $Pr(\bar{\theta}) = 1 - Pr(\underline{\theta})$. D knows nature's selection θ, but A does not.

2. State A selects rising technology level $t \in \mathcal{T} = \mathbb{R}_{\geq 0}$.

3. State 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 of hassling $h \in \mathcal{H} = \mathbb{R}_{\geq 0}$ (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 $U_A = P(0,0) - \kappa_A$ and $U_D = 1 - P(0,0) - \kappa_D$ for States A and D (respectively; I characterize the P function below). When D does not go to war, the game moves to the next stage.

4. State D offers State A some value $x \in [0, 1]$.

5. State A can declare "war" by setting $w_A = 1$ or can "accept" the offer by setting $w_A = 0$. When State A sets $w_A = 1$, State A receives their its updated wartime payoff $U_A = P(t,h) - \kappa_A$, and State D receives $U_D = 1 - P(t,h) - \kappa_D - \frac{h^2}{F(\alpha,\theta)}$, which is their its updated wartime payoff $(1 - P(t,h) - \kappa_D)$ minus their its costs from hassling $(\frac{h^2}{F(\alpha,\theta)})$. When State A sets $w_A = 0$, State A receives payoff $U_A = x$ and State D

receives $U_D = 1 - x - \frac{h^2}{F(\alpha,\theta)}$.

For ease, I summarize all possible game outcomes in Table 1.

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

Table 1: Summarized payoffs for actors.

The function $P : \mathcal{T} \times \mathcal{H} \to [0,1]$ is the likelihood that A wins in a war. I assume functional form $P(t,h) = max\{min\{1, \rho + t - h\}, \rho\}$ with $\rho \in [0,1]$, which implies P is weakly increasing in t and $-h$, and that P falls between ρ and 1 inclusive. The constant ρ is A's likelihood of winning a war before the rising technology or hassling comes to fruition (where $t = 0$ and $h = 0$). That $P(t,h)$ must be (weakly) greater than ρ implies that while hassling can degrade investments in rising technology t, hassling can never make A a declining state, and, at most, hassling returns A to a baseline war victory likelihood of ρ.[14] This functional form may be perceived as undesirable as it contains "kinks" (at ρ and 1) that could drive results; to address this issue, I will demonstrate cases, analysis, and alternate functional forms below and in the Appendix where the kinks play no role in any actor's behavior.

When D initiates war in Stage 3, the actors fight over the asset of normalized value 1. This stage closely resembles the Fearon (1995) costly lottery treatment of war. A's likelihood of winning in war is $P(0,0) = \rho$, and $\kappa_A > 0$ and $\kappa_D > 0$ are A's and D's costs from war, respectively.

[14] While low-level conflict could potentially turn a state into a declining power, this falls outside of the scope of the analysis.

When A initiates war in Stage 5, the fighting happens after the investments in rising technology and hassling have both occured. The P function, κ_A, and κ_D have been discussed above. The expression $\frac{h^2}{F(\alpha,\theta)}$ denotes the additional costs that D incurs from hassling, where the function $F : \{\mathcal{A} \times \Theta\} \rightarrow \{\mathbb{R}_{>0}\}$, and where F is 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 A's investment). Throughout the paper, I will define parameters $\underline{\alpha} \in \mathcal{A}$ and $\bar{\alpha} \in \mathcal{A}$ with $\underline{\alpha} < \bar{\alpha}$. For ease, I assume that A faces no costs to invest in the rising technology t. I explore a model with these costs in an extension in the Appendix, and the substantive results do not change.

When A accepts in Stage 5, this represents a bargained outcome after the rising technology and hassling have both occurred. The value x denotes the offer D makes to A, and D still incurs the costs from hassling.

2.2 Equilibrium Concepts and Assumptions

I limit attention to pure strategy perfect Bayesian Nash Equilibria.[15] The actions taken in the game depend on the public and private components of D's capabilities. A strategy for State D is a mapping from the selected level of rising technology t and its hassling capabilities to its action space (consisting of w_D, h, and x), or $\sigma_D : (\mathcal{T}, \mathcal{A}, \Theta) \rightarrow \{0,1\} \times \mathcal{H} \times [0,1]$. Because State A does not know the value of θ, A's strategy is a mapping from the known parameter α to its action space (consisting of t and w_A), or $\sigma_A : \mathcal{A} \rightarrow \mathcal{T} \times \{0,1\}$. I let σ denote a pair of strategies or $\sigma = (\sigma_A, \sigma_D)$, and a strategy profile $\sigma^* = (\sigma_A, \sigma_D)$ constitutes a Bayesian equilibrium if $\sigma_D(\mathcal{T}, \alpha, \theta)$ is a best response to σ_A, and $\sigma_A(\alpha)$ is a best response to σ_D based on the known capabilities parameter and expectations over D's type. For ease, I limit myself to pure strategy equilibria, I let x^*, h^*, w_A^*, w_d^*, and t^* denote equilibrium values of the various actions, and I let $\sigma^*(\alpha, \theta)$ denote the equilibrium for parameter $\alpha \in \mathcal{A}$

[15] Considering mixed strategies does not change the substance of the results. In short, mixed strategies would lead to multiple equilibria at parameters where players are indifferent between actions.

and type $\theta \in \Theta$.

At this point, I can formally define a "deterrence failure." The logic behind the definition is that D wants to possess a set of technological capabilities that produces good outcomes (i.e., greater utility); thus, if improvements in hassling capabilities produce worse outcomes, it implies that the increased capability is prompting A to undertake problematic behavior.

Definition*: Improvements in publicly observed hassling capabilities (i.e. moving from* $\underline{\alpha}$ *to* $\bar{\alpha}$ *with* $\underline{\alpha} < \bar{\alpha}$*) produces a* ***deterrence failure*** *when,* $U_D\left(\sigma^*(\theta, \underline{\alpha})\right) \geq U_D\left(\sigma^*(\theta, \bar{\alpha})\right)$ *for all* $\theta \in \Theta$ *and* $U_D\left(\sigma^*(\theta, \underline{\alpha})\right) > U_D\left(\sigma^*(\theta, \bar{\alpha})\right)$ *for some* $\theta \in \Theta$.

With this definition, a deterrence failure happens only when every possible type of D experiences the same or less expected utility when its public capabilities are high than when its public capabilities are low.[16]

As a final assumption, I will limit my analysis below to scenarios where the constraints on the P function and x do not bind. These technical conditions have the result that in equilibrium, $P \in (p, 1)$ and $x \in (0, 1)$. This assumption is useful because it eliminates the possibility that the kink in the P function or the requirement that $x \in [0, 1]$ drive any of the results.[17] Furthermore, this allows the analysis to be straightforward, without needing to consider excessive casework (i.e., defining where the constraints do or do not bind) while still using a benign functional form.[18] In the Appendix I include alternate functional forms and

[16] I discuss why I choose this definition rather than using D's ex-ante expected utility in the Appendix. To briefly summarize, when using the ex-ante expected utility definition of a deterrence failure, some types of D could attain strictly greater utilities when a deterrence failure occurs. Thus, the expected utility definition is problematic because if D had more information than A about it's type, some types of D would actually prefer to experience a deterrence failure.

[17] This game is different from typical multi-stage games illustrating commitment problems, where the inability of the rising state to make a large enough concession in the first round produces a preventive war (Fearon, 1995).

[18] This assumption allows me to define $x^* = \rho + t^* - h^* - \kappa_A$. Without this assumption, I must consider $x^* = min\left\{max\left\{min\left\{max\left\{\rho + t^* - h^*, \rho\right\}, 1\right\} - \kappa_A, 0\right\}, 1\right\}$ as part of A's optimization.

specifications that do not rely on this assumption. The findings are substantively identical but the analysis is more complicated.

3 Equilibrium

The intuition for the game's equilibrium is as follows. In the fourth and fifth stages, D does strictly better always avoiding war, which D can always accomplish by making A an offer equal to its wartime utility $x^* = \rho + t^* - h^* - \kappa_A$. In the third stage, D reacts to A's selected level of rising technology (t^*) by either going to war (setting $w_D^* = 1$) or not going to war (setting $w_D^* = 0$) and selecting an optimized hassling level h^* to counter the selected t^*. In the second stage, A selects an optimal level of t^* based on how A expects D will respond to that t^*, which is conditional on D's known parameter (α) and the values and distribution of D's type (θ).

Because A faces no costs to investing in the rising technology, A wants to invest in high levels, but it knows that if it selects too great a t, D will respond with war. Thus, A's expected utility from increasing its investment is increasing, unless it provokes D to go to war. D's cutoff strategy is determined by D's hassling capabilities. For brevity, I will use the notation $D(\alpha, \theta)$ to refer to a State D that has type θ and parameter α. Because α is known and because A knows D is either type $\underline{\theta}$ or $\bar{\theta}$, for a given $\alpha \in \{\underline{\alpha}, \bar{\alpha}\}$, A will select either a t that would make a $D(\alpha, \underline{\theta})$ indifferent between war and hassling, or select a t that would make a $D(\alpha, \bar{\theta})$ indifferent between war and hassling. I let $t(\alpha, \theta)$ denote the level of rising technology that would make a $D(\alpha, \theta)$ indifferent between war and hassling. If A selects $t(\alpha, \underline{\theta})$, A avoids war altogether. If A selects $t(\alpha, \bar{\theta})$, $D(\alpha, \underline{\theta})$ (but not $D(\alpha, \bar{\theta})$) will respond with war.

Thus the equilibrium takes one of two cases: a case where A optimally avoids war and a case where A sometimes risks war. Proposition 1 lists some features of the equilibrium, and the Appendix contains the full equilibrium.

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* $Q(\alpha) = Pr(\underline{\theta})\,(\kappa_A + \kappa_D) + \big(Pr(\bar{\theta}) - Pr(\underline{\theta})\big)\,\frac{F(\alpha,\underline{\theta})}{4} - Pr(\bar{\theta})\frac{F(\alpha,\bar{\theta})}{4}$*:*

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

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

With $Q(\alpha) = Pr(\underline{\theta})\,(\kappa_A + \kappa_D) + \big(Pr(\bar{\theta}) - Pr(\underline{\theta})\big)\,\frac{F(\alpha,\underline{\theta})}{4} - Pr(\bar{\theta})\frac{F(\alpha,\bar{\theta})}{4}$.

Proof: All proofs are in the Appendix.

In Case 1, A will seek to avoid war altogether by selecting a level of rising technology 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 $1 - \rho - \kappa_D$. Because A 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, A will sometimes risk war by selecting a level of rising technology 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 $1 - \rho - \kappa_D$. Because A 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. Thus,

$Q(\alpha)$ is simply the cutpoint where, for values below the cutpoint, A does better by selecting the higher t^*—that is, selecting $t(\alpha, \bar{\theta})$—and sometimes risking war.

4 Results

In this section, I will first discuss how improvements in public hassling capabilities can lead to D becoming more predictable or A becoming emboldened, thus producing a deterrence failure. I then discuss how these two mechanisms are the *only* way a deterrence failure can occur in the model. In Section 6, I describe what happens when private hassling capabilities improve, the relationship between deterrence failure and states' welfare, and what happens when improvements in hassling capabilities also change wartime capabilities. I include more details on the mechanisms in the Appendix.

4.1 Predictability and Deterrence Failure

In the game, State D (weakly) benefits from its private information.[19] An improvement in public hassling capabilities makes State D more predictable when $D(\bar{\alpha}, \underline{\theta})$ and $D(\bar{\alpha}, \bar{\theta})$ select more similar levels of hassling than $D(\underline{\alpha}, \underline{\theta})$ and $D(\underline{\alpha}, \bar{\theta})$. As a result, when public capabilities are higher, State A can select an investment in rising technology that better extracts the benefit that D attains from its private information. This occurs despite the likelihood of war remaining unchanged. Figure 1 visualizes this intuition using the results from a fully parameterized example. For the selected parameters, under both $\underline{\alpha}$ and $\bar{\alpha}$, the equilibrium is described in Case 1 in Proposition 1, where A always avoids war.

Under the selected parameters, A optimally selects the largest investment in rising technology t that will keep D from ever declaring war. Because A observes D's parameter α but not D's type θ, the largest t that A can select while still avoiding war is the t that would keep $D(\alpha, \underline{\theta})$

[19] In an equivalent setup but with full information, A would always select the level of rising technology that makes D indifferent between war and hassling, thus always giving D its wartime utility regardless of its type.

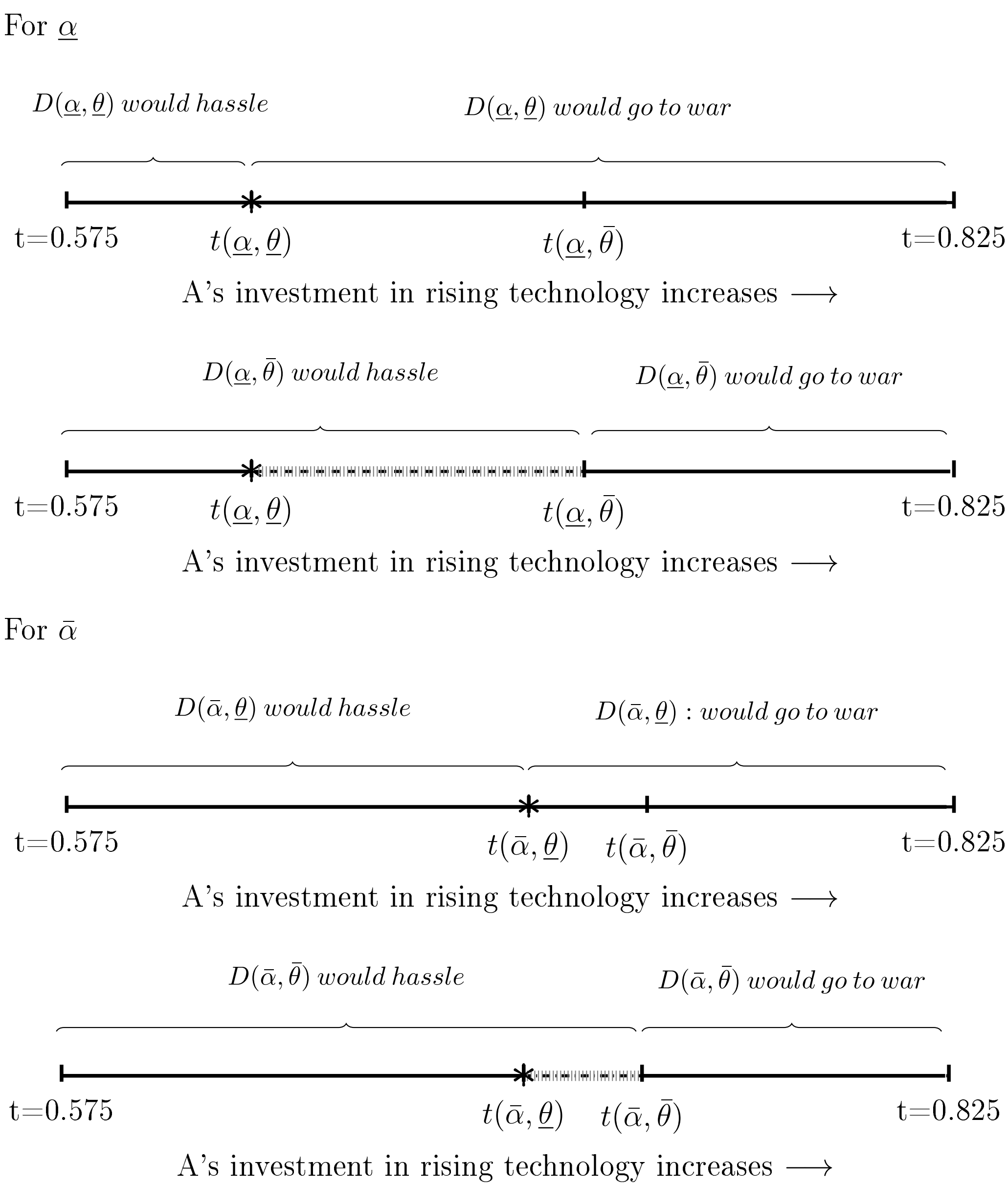

Figure 1: Optimal rising technology levels and D's response (Predictability).

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

indifferent between hassling and going to war, or $t(\alpha, \underline{\theta})$. In the top two panels of Figure 1 (where $\alpha = \underline{\alpha}$), A's optimal investment level $t^* = t(\underline{\alpha}, \underline{\theta})$ is indicated by asterisks. In the top panel, the selected t^* makes $D(\underline{\alpha}, \underline{\theta})$ indifferent between war and hassling, implying that it will attain its wartime utility. In the second panel, the space between the selected level of rising technology and the level of rising technology 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 $D(\underline{\alpha}, \bar{\theta})$ is able to keep its type private in the lead-up to A's selection of t by successfully posturing,[20] it can attain some bargaining surplus.

However, when D's public hassling capabilities improve, D's private information can become less valuable. Because A has some uncertainty over how willing D is to hassle, A must benchmark its investment against type $\underline{\theta}$ if it seeks to avoid war. If improvements in D's known hassling abilities reduce the importance of D's unknown type, then A does not need to scale back as much to prevent war. Comparing the top two and bottom two panels of Figure 1 illustrates how this may occur. In the top two panels, when α is low, there is a significant gap between A's selected investment $t(\underline{\alpha}, \underline{\theta})$ and the point that makes $D(\underline{\alpha}, \bar{\theta})$ indifferent between war and hassling ($t(\underline{\alpha}, \bar{\theta})$). When α increases, there is now a much smaller gap between the amount of technology A selects and the amount that makes $D(\bar{\alpha}, \bar{\theta})$ indifferent between hassling and war. This implies that as D becomes publicly better at hassling (moving from $\underline{\alpha}$ to $\bar{\alpha}$), A is able to select a level of investment in rising technology closer to the point that gives $D(\bar{\alpha}, \bar{\theta})$ its wartime payoff. Put another way, in the bottom two panels, when $\alpha = \bar{\alpha}$, 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 has made D's private type a smaller factor in determining D's selected hassling levels, and therefore has made D more predictable. This case defines a deterrence failure for D because an increase in public

[20] The "posture" can be thought of as $D(\underline{\alpha}, \bar{\theta})$ telling A that it is actually a $D(\underline{\alpha}, \underline{\theta})$, and that, in order to avoid war, A must select a low level of rising technology.

hassling capabilities means A can select greater and better tailored levels of rising technology.

For the hassling game characterized above, Proposition 2 defines the necessary conditions under which an improvement in public hassling capabilities makes D more predictable, producing a deterrence failure.

Proposition 2 (Predictability): *Under the Predictability Conditions, A 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 Predictability Conditions hold, then improvements from $\underline{\alpha}$ to $\bar{\alpha}$ produce a deterrence failure.*

Proposition 2 generalizes the intuition in the numerical example above. The first two Predictability Conditions imply that across parameters $\underline{\alpha}$ and $\bar{\alpha}$, the behavior of A and D is characterized in Case 1 of Proposition 1. The third condition captures the differences in how types $\underline{\theta}$ and $\bar{\theta}$ play the game across parameters $\underline{\alpha}$ and $\bar{\alpha}$. Intuitively, for a given α, A selects a fixed t^* 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 $h^*(\theta) = \frac{F(\alpha,\theta)}{2}$ and attains a greater utility, captured in the $\frac{F(\alpha,\bar{\theta})}{4} - \frac{F(\alpha,\underline{\theta})}{4}$ term in $D(\alpha, \bar{\theta})$'s utility function. When $F(\alpha, \bar{\theta}) - F(\alpha, \underline{\theta})$ is small, it implies that $D(\alpha, \underline{\theta})$ and $D(\alpha, \bar{\theta})$ play the game similarly (the h^* are similar), implying that D's private information is not particularly valuable to D (formally, $\frac{F(\alpha,\bar{\theta})}{4} - \frac{F(\alpha,\underline{\theta})}{4}$ is small). Therefore, when the third condition holds, the improvement in α degrades the value of D's private information and allows A's selected t^* to be better calibrated, thus producing the deterrence failure.

4.2 Emboldening and Deterrence Failure

Recall that in the game, A wants to invest in the rising technology as much as possible, but may be deterred from investing if further investment leads to a greater probability of war. Whether A is willing to invest to a level that risks war is contingent upon the benefit that A receives from the investment, which is dependent upon D's selected hassling levels. When D's public hassling capabilities shift from low ($\underline{\alpha}$) to high ($\bar{\alpha}$), D is both more willing to hassle over a wider range of rising technologies and will hassle more for a given investment level t. If A is emboldened, D's shift in capabilities incentivizes A to invest more aggressively, knowing that, while D may initiate war, A holds a better bargaining position following D's more aggressive hassling when D does not initiate war. Put another way, an improvement in public hassling capabilities emboldens A when the benefit of investing more aggressively now outweighs the greater risk of war.[21]

Figure 2 visualizes this intuition using the results from a fully parameterized example. As can be seen, A's expected utility is increasing in the selected t (moving right along the x-axis), until this selection goes too far and 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 $t(\alpha, \bar{\theta})$, producing the second discontinuity.

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, A optimally selects $t^* = t(\underline{\alpha}, \underline{\theta})$, which is characterized by the asterisk. In response, both $D(\underline{\alpha}, \underline{\theta})$ and $D(\underline{\alpha}, \bar{\theta})$ will always hassle. It is worth highlighting what A does not do when $\alpha = \underline{\alpha}$: A does not select $t(\underline{\alpha}, \bar{\theta})$.[22] This "move not taken" captures A's trade-off between a greater investment in its rising technology

[21] Spaniel and Malone (2019) find a similar mechanism can arise following improved economic interdependence.

[22] Note that A would never select a t greater than the value that would make $D(\underline{\alpha}, \bar{\theta})$ indifferent between hassling and war because this would always result in war, which is strictly worse for A.

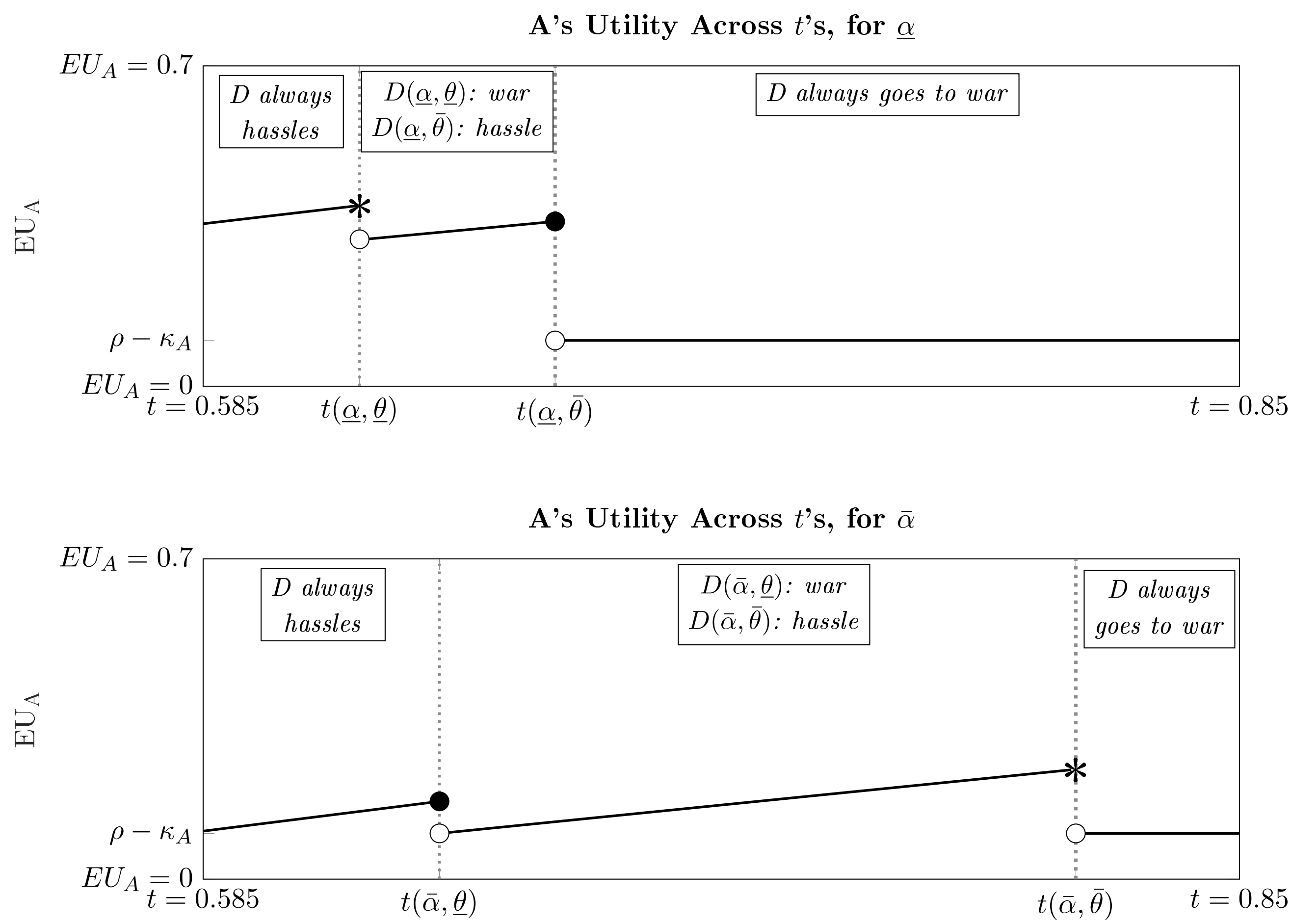

Figure 2: Optimal rising technology levels and A's utility (Emboldening).

A's selected level of investment under parameters $\underline{\alpha}$ and $\bar{\alpha}$ are denoted by the asterisks. D's response to the selected t's are labeled. The parameters are $\kappa_D = 0.1$, $\kappa_A = 0.4$, $\rho = 0.5$, $Pr(\underline{\theta}) = 0.2$, $Pr(\bar{\theta}) = 0.8$, $F(\underline{\alpha}, \underline{\theta}) = 0.5$, $F(\underline{\alpha}, \bar{\theta}) = 0.7$, $F(\bar{\alpha}, \underline{\theta}) = 0.6$, $F(\bar{\alpha}, \bar{\theta}) = 1.3$.

and a higher risk of war. The upside to selecting $t(\underline{\alpha}, \bar{\theta})$, is that A has an advantage when facing $D(\underline{\alpha}, \bar{\theta})$, thus giving A a larger final utility.[23] However, the downside for A of selecting a $t(\underline{\alpha}, \bar{\theta})$ is that now $D(\underline{\alpha}, \underline{\theta})$ will go to war, increasing the likelihood of war from zero to the likelihood that D is type $\underline{\theta}$. The difference between $t(\underline{\alpha}, \underline{\theta})$ and $t(\underline{\alpha}, \bar{\theta})$ can be thought of as the amount of rising technology that A could invest in if A were willing to go to war with $D(\underline{\alpha}, \underline{\theta})$. When $\alpha = \underline{\alpha}$, A's upside to doing better against $D(\underline{\alpha}, \bar{\theta})$ does not outweigh the costs of increasing the likelihood of war against $D(\underline{\alpha}, \underline{\theta})$. In the figure, this is shown by the

[23]For the selected parameters, when A selects $t = t(\underline{\alpha}, \underline{\theta}) = 0.625$, A attains utility 0.375 when nature sets $\theta = \bar{\theta}$. When A selects $t = t(\underline{\alpha}, \bar{\theta}) = 0.675$, A attains utility 0.425 when nature sets $\theta = \bar{\theta}$.

asterisk at $t(\underline{\alpha}, \underline{\theta})$ giving A more utility than the point at $t(\underline{\alpha}, \bar{\theta})$.

When $\alpha = \bar{\alpha}$, where the gameplay is described in Case 2 in Proposition 1, A faces a new trade-off. Selecting $t(\bar{\alpha}, \bar{\theta})$ will provoke $D(\bar{\alpha}, \underline{\theta})$ to go to war. However, as the bottom panel illustrates, A now faces a much greater upside to selecting $t(\bar{\alpha}, \bar{\theta})$ over $t(\bar{\alpha}, \underline{\theta})$ because these cutpoints are so different, representing a large opening for investing in rising technology if A is willing to risk some likelihood of war.[24] Now, A's upside to doing better against $D(\bar{\alpha}, \bar{\theta})$ outweighs the increased likelihood of war with $D(\bar{\alpha}, \underline{\theta})$, as shown by the asterisk at $t(\bar{\alpha}, \bar{\theta})$ giving A more utility than the point at $t(\bar{\alpha}, \underline{\theta})$. Intuitively, this case defines a deterrence failure for D because an increase in public hassling capabilities emboldened A to pursue a riskier investment strategy in rising technology and sometimes provoke war, both to D's detriment.

Proposition 3 defines, for the hassling game characterized above, the necessary conditions for improvements in public hassling capabilities to embolden A and produce a deterrence failure.

Proposition 3 (Emboldening): *Under the Emboldening Conditions, A 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 conditions hold, then improvements from $\underline{\alpha}$ to $\bar{\alpha}$ produce a deterrence failure.*

Proposition 3 generalizes the intuition in the numerical example above. The $Q(\underline{\alpha}) \geq 0$ condition implies that A will not risk war and so will select $t^* = t(\underline{\alpha}, \underline{\theta})$, which will give $D(\underline{\alpha}, \bar{\theta})$ a utility above its wartime payoff. $D(\underline{\alpha}, \underline{\theta})$, meanwhile, is indifferent between war and hassling so it gets its wartime utility. The $Q(\bar{\alpha}) < 0$ condition implies that A will risk war with $D(\bar{\alpha}, \underline{\theta})$ by selecting $t^* = t(\bar{\alpha}, \bar{\theta})$, which will result in both $D(\bar{\alpha}, \underline{\theta})$ and $D(\bar{\alpha}, \bar{\theta})$

[24] For the selected parameters, when A selects $t = t(\bar{\alpha}, \underline{\theta}) = 0.65$, A attains utility 0.1 when nature sets $\theta = \bar{\theta}$. When A selects $t = t(\bar{\alpha}, \bar{\theta}) = 0.825$, A attains utility 0.275 when nature sets $\theta = \bar{\theta}$.

attaining their wartime utilities. Because $D(\bar{\alpha}, \bar{\theta})$ does worse than $D(\underline{\alpha}, \bar{\theta})$, this produces a deterrence failure for D.

4.3 When (and Only When) Improved Capabilities Create Deterrence Failures

With the model and assumptions characterized above, a deterrence failure for D is possible *only* when the Predictability Conditions or the Emboldening Conditions hold.

Proposition 4:*Improvements in* α *produce a deterrence failure if and only if the Emboldening or Predictability Conditions hold.*

Based on Proposition 4, I have comprehensively characterized when improved public hassling capabilities produce worse outcomes for the state making these improvements. This result is important for two reasons. First, understanding how improvements in military capabilities can inadvertently lead to detrimental strategic responses in rivals and more conflict is an important topic. And, this paper uses a fairly flexible model to address this topic through the context of a fundamental strategic interaction. By identifying the *only* two ways improvements in military capabilities can lead to worse outcomes in this environment, I have characterized two critical mechanisms for deterrence failure. Second, while I do not claim to have identified every possible kind of deterrence failure in every possible model, this result suggests that looking for other sources of deterrence failures either must consider a different strategic interaction, or will rely on modify the structure of the game beyond anything covered here or in Section 6.

5 Discussion of Cases

In this section, I describe how the model offers insights into the 2003 U.S. invasion of Iraq and specifically why Saddam Hussein made the choices he did in the months before the war began. The model also helps to explain the mechanism that produced the "stability-instability paradox" during the Cold War, when the "stability" of U.S.-Soviet relations at the nuclear level was paired with "unstable" low-level conflict between the superpowers.

5.1 Emboldening: Saddam's Gamble

An important empirical implication of the theory offered here is that when a defending state becomes better at hassling, the rising state may be emboldened to invest in a high level of rising technology, which can lead to an escalation to war. Put another way, improvements in the technological capabilities for low-level conflict (hassling) can lead to more instances of high-level conflict (war).[25]

To summarize some details of the case, within the ceasefire terms of the 1991 Persian Gulf War and following United Nations Security Council Resolutions, the United Nations outlined that Iraq was forbidden from possessing nuclear, biological, or chemical weapons (WMDs hereafter), and was subject to inspections and protocols from a United Nations Special Commission and the International Atomic Energy Agency. Between 1991-1998, Iraq repeatedly violated the terms of the resolutions by engaging in the following behavior: firing warning shots at inspectors to prevent them from inspecting suspected vehicles (1991); preventing inspectors from searching government buildings (1992); submitting incomplete disclosures of weapons programs (1995-1996); and eventually terminating all cooperation with inspectors (1998). These actions (and others) led U.S. and British forces to engage in Operation Desert Fox, a multi-day bombing strike against Iraqi weapons targets in late 1998. Between

[25] This result arises without assumptions that low-level conflict can probabilistically escalate, like that in Powell (2015).

1999-2002, Iraq remained committed to keeping weapons inspectors out, despite intensifying rhetoric within the U.S. over the status of Iraqi WMD programs. And, while in late 2002 Iraq agreed to resume inspections, after over a decade of friction with inspectors, Iraq's change was insufficient, and the United States invaded Iraq in March 2003. In 2004 the Iraq Survey Group released what became known as the "Duelfer Report," that found 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 U.S. invasion of Iraq, three questions must be addressed: Was Iraq "investing in rising technology" as defined above? Did the U.S. demonstrate a robust hassling capability? And, did U.S. hassling capabilities shape Iraqi behavior?

An "investment in rising technology" occurs when a rising state pursues opportunities that strengthen its future military capabilities; after the rising state invests, the defending state observes the investment, may become concerned with their own political future if the investment is allowed to come to fruition, and may hassle or declare war. Within this case, Saddam's investment in rising technology was keeping the weapons inspectors out. Keeping weapons inspectors out strengthened Iraq's future military capabilities. As discussed in Koblentz (2018) and Coe and Vaynman (2020), states typically dislike allowing external observers to inspect their military capabilities because this information could eventually be used against them by adversaries. Thus, in breaking the terms of the United Nations resolutions, Saddam would be in a better military position should a conflict arise. And, keeping weapons inspectors out led the United States to grow concerned that Saddam was developing WMDs and was therefore a rising state. 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.[26]

Of course, I do not want to gloss over the fact that Saddam was not actually building WMD, and that if the U.S. had all the correct information, then the U.S. plausibly would not have invaded Iraq. However, the reason the U.S. did not know the extent of Iraq's WMD programs was because Saddam was keeping weapons inspectors out. A key feature of Saddam's decision to keep weapons inspectors out was that it created misconceptions and uncertainty around Saddam's WMD programs that the U.S. grappled with in the lead up to the 2003 invasion. Ultimately, Saddam did not know the U.S.'s true willingness to accept uncertainty over Iraq's WMD programs, and Saddam created more uncertainty than the U.S. was willing to resolve with hassling.

The second question to answer is whether the U.S. had publicly demonstrated a robust hassling capability. This was indeed the case. In the aftermath of the Cold War, the U.S. successfully transformed much of its technical know-how in building and deploying missiles and bombers for nuclear strikes into precision strike capabilities (U.S. Air Force, 2019a,c,b). Additionally, the efficacy of these capabilities and U.S. public willingness to use them was revealed in a series of conflicts, including the NATO bombing of Bosnia and Herzegovina, the NATO bombing of Yugoslavia, and U.S. strikes against Iraq in 1998 in Operation Desert Fox.[27]

Third, there is ample evidence that Iraq was aware of U.S. hassling capabilities and that Saddam's decision to refuse entry to the weapons inspectors was influenced by these capabilities. After Saddam was captured, in interviews, he was asked why he turned away weapons inspectors, which resulted in the following retelling (Battle, 2009):

[26] Saddam's internal deliberations are covered in Woods *et al.* (2006, 15-16, 30, 91, 96-97, 125) and Woods *et al.* (2011, 257-258).

[27] The aggressive sanctions against Saddam and the support for third party actors in Iraq also represent U.S. hassling capabilities, as discussed in McCormack and Pascoe (2017).

> *Even though Saddam claimed Iraq did not have WMD, the threat from Iran was the major factor as to why he did not allow the return of the UN inspectors. Saddam stated he was more concerned about Iran discovering Iraq's weaknesses and vulnerabilities than the repercussions of the United States for his refusal to allow UN inspectors back into Iraq. [...] Saddam indicated he was angered when the United States struck Iraq in 1998. Saddam stated Iraq could have absorbed another United States strike for he viewed this as less of a threat than exposing themselves to Iran.*

This quote suggests that Saddam understood that refusing entry to the weapons inspectors might be met by a forceful U.S. response, but he believed that response would take the form of Operation Desert Fox—hassling, not war. And while Saddam may have had incentive to lie, in internal discussions leading up to the 2003 invasion, Saddam commented frequently on the U.S. reliance on air strikes—not ground invasion—for international conflict, a sentiment echoed by Saddam's advisers (Woods *et al.*, 2006, 15-16, 30, 96-97, 125).

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

That Saddam changed his behavior in response to improved U.S. hassling capabilities (i.e., improvements that facilitate a low level of 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. Somewhat similar results exist in the literature—both Bas and Coe (2016) or Joseph (2020) suggest that being better at low-level operations like preventive strikes can lead to more low-level operations—but the literature cannot currently explain how being better at hassling can lead to an escalation to war. For example, the Joseph result suggests that being better at hassling can lead to more weapons proliferation and more hassling occurring in the world, but never finds that it can produce more war.[28]

The theory and case above suggest a new perspective on the 1990s "revolution in military affairs," which treated developments in precision strike and cyber capabilities as fundamentally altering the way conflict— especially low-level conflict— would be fought. This paper 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 U.S.'s increased reliance on hassling technologies and was emboldened, making a gamble in turning away weapons inspectors, believing that his decision would likely be met with a limited response. This case suggests that improvements in hassling capabilities can shift the strategic decisions that revisionist states make, and can ultimately result in devastating and destabilizing wars.

5.2 Predictability and the Stability-Instability Paradox

When the predictability mechanism is in play, rising states invest more in the rising technology than they otherwise would because they recognize that the more aggressive investments will be met with hassling rather than war. This suggests that if the defending state were worse at hassling, the threat of escalation to war would be more credible for a wider range of investments, and war could better function as a deterrent threat.

This logic helps to explain a phenomenon that scholars observed during the Cold War:

[28] Bas and Coe (2016) find substantively similar results, where being better at preventive attacks, through strategic affects, can make preventive attacks more likely.

the "stability-instability paradox" (Snyder, 1965; Jervis, 1984). To elaborate on the paradox, while mutually assured destruction prevented the United States and Soviet Union from engaging in nuclear war (stability), the superpowers did compete militarily for influence in numerous conflicts around the globe (instability). Indeed, during the Cold War, U.S. and U.S.S.R. regularly engaged in territorial power grabs (e.g., the Soviet invasion of Afghanistan) or provided pro-communist or pro-western rebels and governments with weapons, resources, and cash (e.g., the U.S. provided arms, and training to Guatemalan rebels). And, these behaviors fall within the scope of what is analyzed here. For example, the U.S.S.R. invaded Afghanistan in 1979 to ensure the Afghan government remained a client of the Soviet Union and thus a source of future geopolitical strength, making the invasion a form of "investing in rising technology." In response, the U.S. did not go to war; rather, in the initial response, the U.S. attempted to degrade the Soviet investment by supporting mujahideen rebels, making its response a form of "hassling."[29]

Interestingly, in Snyder's 1965 initial presentation of the paradox, he leaves open the possibility for a different outcome at the nuclear (or as he puts, "strategic") level. In the initial text, Snyder states the following:

> *Thus firm stability in the strategic balance tends to destabilize the conventional balance and also to activate the lesser nuclear "links" between the latter and the former. But one could argue precisely the opposite—that the greater likelihood of gradual escalation due to a stable strategic equilibrium tends to deter both conventional provocation and tactical nuclear strikes – thus stabilizing the overall balance. The first hypothesis probably is dominant, but it must be heavily qualified by the second, since nations probably fear the possibility of escalation "all the way" nearly as much as they fear the possibility of an "all-out" first strike.*

[29] The initial aim of U.S. efforts in Afghanistan was to weaken the U.S.S.R.; only later, after the mujahideen proved capable, did the U.S. aim to overthrow the Soviet backed government (Coll, 2005).

This paper can speak to Snyder's initial uncertainty over the existence of the paradox by presenting a critical scope condition: instability at lower levels is exacerbated when each state possesses a capacity to engage at these lower levels, thus reducing the risk of escalation and creating high-level stability. The predictability mechanism demonstrates how a state being highly capable at low-level conflict (hassling) can lead a rival to make more, better calibrated investments in rising technology, which means the first state has to engage in more hassling—with the result that the hassling state experiences worse outcomes. Thus, the theory here suggests that it was not simply stability at the nuclear level that created an abundance of conflict between the superpowers at lower levels, but rather that U.S. and U.S.S.R. investments in abilities to conduct low-level operations mattered too. Both the CIA and the KGB embraced a wide portfolio of covert operations as a means for projecting and securing their government's influence (O'Brien, 1995). And, within each state, each agency took on a more pronounced role in strategic operations: for example, within the United States, funding for the newly formed CIA grew rapidly after 1945 (Gaddis, 2005; Snider, 2015). The theory here suggests that by becoming better at low-level operations, the U.S. and U.S.S.R. opened the door to more aggressive power grabs by the other because both could predict that their attempts would, at worst, be met with a low-level response.

Today, several scholars and practitioners have observed a new emergence of great power competition and a new stability-instability paradox (Lindsay *et al.*, 2016; DoD, 2018). What is unique about this period's paradox is that the expansion in low-level engagement across the globe is conducted not only by states who have established a credible nuclear (or conventional conflict) deterrence, but also involves unbalanced dyads like the U.S. and Iran. The theory here suggests that this new competition could be fueled by the proliferation of precision, drone, and cyber technologies, which make hassling cheap and easy. The prevalence of these technologies in low-level conflicts today offers some support for this theory. To this end, in the Appendix, I discuss what these results mean for best-practices in developing hassling capabilities. The key take-away is that improved hassling capabilities should be accompanied by improved capabilities at higher levels of escalation to successfully prevent deterrence failures.

6 Additional Modeling Results and Extensions

6.1 Additional Results

I include a formal discussion of each of these results in the Appendix.

Can improvements in *private* hassling capabilities (θ) ever produce worse outcomes for D? I find that if D experiences an improvement in private hassling capabilities (i.e. moving from $\underline{\theta}$ to $\bar{\theta}$), then D always attains a weakly greater utility. This finding suggests that improvements in private capabilities are theoretically distinct from improvements in public capabilities, as only improvements in public capabilities can produce deterrence failures.

When deterrence failures occur, what can be said about aggregate welfare? I find that whenever improvements in publicly observed hassling capabilities produce a deterrence failure, then theses improvements also cause a welfare loss. This finding adds nuance to the concept of a deterrence failure. When a deterrence failure occurs following improvements in public hassling capabilities, it is not just that utility is transferred from D to A. Rather, when the Emboldening or Predictability Conditions hold, the combined utility of D and A declines. And, this is not to say that improvements in hassling capabilities always produce a welfare loss: in the Appendix, I show that, under some conditions, improvements in public hassling capabilities can produce a welfare improvement.

What if improved hassling capabilities affect wartime capabilities? I find that if improvements in publicly observed hassling capabilities also improve payoffs from war, then the improvements in public hassling capabilities can never produce a deterrence failure. In the two cases in Proposition 1, type $\underline{\theta}$ D 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 $D(\underline{\alpha}, \underline{\theta})$ following the change, which undermines the conditions for deterrence failure. Whether or not improved hassling capabilities affects war capabilities is an empirical question. For example, if increased investments in precision strike capabilities shift funds in such a way that it undermines the army's ability to conduct a counterinsurgency, then improvements in hassling could lead to worse wartime outcomes depending upon what war fighting capabilities were needed. At a minimum, this observation suggests a possibly useful heuristic for how to make investment decisions in hassling capabilities: so long as improvements in hassling also improve wartime outcomes, then these improvements will not produce a deterrence failure.

6.2 Extensions

How general are these results? Are predictability and emboldening the mechanisms that cause a deterrence failure in other models with different assumptions?[30] Overall, I find, across a wide range of alternate assumptions and models, that deterrence failure arises only if A becomes emboldened or D becomes more predictable. I describe my results below and present the full results in the Appendix.

In this paper's model, A does not incur costs from investing in the rising technology. In Section A7, I examine a model where A does incur costs from investment. I find that deterrence failure only arises from D becoming more predictable or A becoming emboldened.

In my model, I also assume that in equilibrium the final realized P function and offer x have the properties $P \in (\rho, 1)$ and $x^* \in (0, 1)$. In Section A8, I examine a model that utilizes

[30] Generally, when D becomes more predictable, it implies that high-types ($\theta > \underline{\theta}$) play the game more like low-types ($\theta = \underline{\theta}$), with war occurring weakly less. And, when A becomes emboldened, it implies that A selects a level of rising technology resulting in more type D's going to war and with the types of D that do not go to war attain lower utilities.

a different peace and wartime payoff structure where there are no such kinks or bounds and I find substantively similar results. Additionally, these results persist outside of this specific modeling technology, as I show in the general analysis in Section A9.

Additionally, the model here considers only two types. In a model in Section A8 and the general analysis in Section A9, I examine models with a continuum of types, which produces substantively similar results.

Finally, the model here utilizes a specific game structure to bargaining after the rising technology and hassling come to fruition. In Sections A8 and A9, I first consider a more flexible structure that allows for a wide range of bargaining structures, and then use mechanism design to illustrate key intuitions from equilibria with almost no structure to the bargaining protocols. Once again, predictability or emboldening following improvements in public hassling capabilities create deterrence failures.

7 Conclusion

This paper explores a common occurrence in international affairs: one state considers undertaking revisionist activities, knowing that its rivals may respond with a range of policy levers (war or varying degrees of hassling). How do the rival state's capabilities affect how the revisionist state will behave? This scenario applies to a broad range of possible revisionist moves, from building new weapons systems to seizing strategically valuable territory. Despite the frequent occurrence of this scenario, this is the first paper to formally examine it.

To understand this interaction, I present a new game theoretic model where a rising power chooses how much to invest in a rising technology, knowing that it faces a defending power who can respond at multiple conflict escalation levels and who has private information about its willingness to engage at different levels. I find that when the defending state has a high level of capability at low-level conflict, two distinct mechanisms can cause worse outcomes for the defending state: the rival state may be emboldened or the defending state may become predictable. These results arise when improved low-level conflict capabilities negatively interact with the defending state's ability to effectively use its private information.

The results here suggest that political scientists and policymakers need to take a harder look at low-level conflict capabilities. As I show here, 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 paper does not suggest that having many policy options is always bad, it does suggest that having more tools can, under some conditions, lead to systemically worse outcomes. More research is needed on this topic.

References

Anderson, J. W. and Khan, K. (1998) Pakistan politicians' rallying cry: 'let them eat grass', *The Washington Post*. 10

Bas, M. A. and Coe, A. J. (2012) Arms diffusion and war, *Journal of Conflict Resolution*, **56**, 651–674. 1.1

Bas, M. A. and Coe, A. J. (2016) A dynamic theory of nuclear proliferation and preventive war, *International Organization*, **70**, 655–685. (document), 1.1, 1.3, 5.1, 28

Battle, J. (2009) *Saddam Hussein Talks to the FBI*, National Security Archive. 5.1

Benson, B. and Smith, B. C. (2020) Do alliances cause commitment problems and war?, *Working Paper*. 2, 1.1

Benson, B. V., Meirowitz, A. and Ramsay, K. W. (2016) Changing capabilities, uncertainty, and the risk of war in crisis bargaining, *Research & Politics*, **3**, 2053168016657687. 1.3

Coe, A. J. (2018) Containing rogues: A theory of asymmetric arming, *The Journal of Politics*, **80**, 000–000. (document), 1.1, 1.3

Coe, A. J. and Vaynman, J. (2020) Why arms control is so rare, *American Political Science Review*, pp. 1–14. (document), 5.1

Coll, S. (2005) *Ghost wars: The secret history of the CIA, Afghanistan, and Bin Laden, from the Soviet invasion to September 10, 2001*, Penguin. 29

Copeland, D. C. (2001) *The origins of major war*, Cornell University Press. 1.1

Debs, A. and Monteiro, N. P. (2014) Known unknowns: power shifts, uncertainty, and war, *International Organization*, **68**, 1–31. (document), 1.1, 1.3

DoD (2018) Summary of the 2018 national defense strategy of the united states of america, Tech. rep., Department of Defense Washington United States. 5.2

Duelfer, C. (2004) *Comprehensive Report of the Special Advisor to the DCI on Iraq's WMD*, vol. 1, Central Intelligence Agency. 5.1

Fearon, J. D. (1995) Rationalist explanations for war, *International organization*, **49**, 379–414. (document), 1.1, 4, 1.3, 2.1, 17

Fearon, J. D. (1996) Bargaining over objects that influence future bargaining power. 1.1

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. 1.1, 1.3

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. 1.1

Gaddis, J. L. (2005) *Strategies of Containment*, Oxford University Press. 5.2

Gartzke, E. and Lindsay, J. R. (2015) Weaving tangled webs: offense, defense, and deception in cyberspace, *Security Studies*, **24**, 316–348. 6

Gartzke, E. and Lindsay, J. R. (2017) Thermonuclear cyberwar, *Journal of cybersecurity*, **3**, 37–48. 1.1, 1.3

Harel, A. and Benn, A. (2018) No longer a secret: How israel destroyed syria's nuclear reactor, *Haaretz*. 1.1

Jervis, R. (1984) The illogic of american nuclear strategy. 5.2

Joseph, M. (2020) The secret failure of nuclear containment, *Working Paper*. (document), 1.1, 1.3, 5.1

Koblentz, G. 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. 5.1

Lanoszka, A. (2016) Russian hybrid warfare and extended deterrence in eastern europe, *International Affairs*, **92**, 175–195. 1.1

Levy, J. S. (1987) Declining power and the preventive motivation for war, *World Politics*, **40**, 82–107. 1.3

Levy, J. S. (2011) Preventive war: Concept and propositions, *International Interactions*, **37**, 87–96. 5, 1.3

Lindsay, J., Gartzke, E., Greenhill, I. K. M. and Krause, P. J. (2016) Coercion through cyberspace: The stability-instability paradox revisited. 5.2

Mazarr, M. J. (2015) Struggle in the gray zone, *War on the Rocks*. 1.1

McCormack, D. and Pascoe, H. (2017) Sanctions and preventive war, *Journal of Conflict Resolution*, **61**, 1711–1739. 3, 1.1, 1.3, 27

Mehta, R. N. and Whitlark, R. E. (2017) The benefits and burdens of nuclear latency, *International Studies Quarterly*, **61**, 517–528. 1.3

Meirowitz, A., Morelli, M., Ramsay, K. W. and Squintani, F. (2019) Dispute resolution institutions and strategic militarization, *Journal of Political Economy*, **127**, 378–418. 1.3

Morrow, J. D. (1989) Capabilities, uncertainty, and resolve: A limited information model of crisis bargaining, *American Journal of Political Science*, pp. 941–972. 1.3

O'Brien, K. A. (1995) Interfering with civil society, *International Journal of Intelligence and Counter Intelligence*, **8**, 431–456. 5.2

Powell, R. (2006) War as a commitment problem, *International organization*, **60**, 169–203. (document), 1.1, 1.3

Powell, R. (2015) Nuclear brinkmanship, limited war, and military power, *International Organization*, **69**, 589–626. (document), 1.3, 25

Ramsay, K. W. (2017) Information, uncertainty, and war, *Annual Review of Political Science*, **20**, 505–527. 11

Reiter, D. (2005) Preventive attacks against nuclear programs and the "success" at osiraq, *Nonproliferation Review*, **12**, 355–371. 1.1

Schelling, T. C. (1980) *The strategy of conflict*, Harvard university press. (document), 1.3

Schram, P. (Forthcomming) Hassling: How states prevent preventive wars, *American Journal of Political Science*. (document), 2, 1.1, 5

Schultz, K. A. (2010) The enforcement problem in coercive bargaining: Interstate conflict over rebel support in civil wars, *International Organization*, **64**, 281–312. 1.3

Slantchev, B. L. and Tarar, A. (2011) Mutual optimism as a rationalist explanation of war, *American Journal of Political Science*, **55**, 135–148. 1.3

Snider, L. B. (2015) *The Agency and the Hill*, Government Printing Office. 5.2

Snyder, G. H. (1965) The balance of power and the balance of terror, in *World in Crisis* (Ed.) F. H. Hartmann, The Macmillan Company, chap. 20, pp. 180–191. 5.2

Spaniel, W. (2019) *Bargaining over the bomb: the successes and failures of nuclear negotiations*, Cambridge University Press. (document), 1.1, 1.3

Spaniel, W. and Bils, P. (2018) Slow to learn: Bargaining, uncertainty, and the calculus of conquest, *Journal of Conflict Resolution*, **62**, 774–796. 1.3

Spaniel, W. and Malone, I. (2019) The uncertainty trade-off: Reexamining opportunity costs and war, *International Studies Quarterly*, **63**, 1025–1034. 21

U.S. Air Force (2019a) Agm-86b/c/d missiles, Tech. rep. 9, 5.1

U.S. Air Force (2019b) B-1b lancer, Tech. rep. 9, 5.1

U.S. Air Force (2019c) B-52h stratofortress, Tech. rep. 9, 5.1

Votel, J. L., Cleveland, C. T., Connett, C. T. and Irwin, W. (2016) Unconventional warfare in the gray zone, *Joint Forces Quarterly*, **80**. 1.1

Woods, K. M., Palkki, D. D. and Stout, M. E. (2011) *The Saddam Tapes: The Inner Workings of a Tyrant's Regime, 1978–2001*, Cambridge University Press. 26

Woods, K. M., Pease, M. R., Stout, M. E., Murray, W. and Lacey, J. G. (2006) Iraqi perspectives project., Tech. rep., United States Joint Forces Command, Norfolk, VA. 26, 5.1