

The Shadow of Deterrence: Why capable actors engage in contests short of war

J Andrés Gannon ^{*1}, Erik Gartzke ², Jon R. Lindsay ³, and Peter Schram ⁴

¹*Stanton Nuclear Security Fellow, Council on Foreign Relations*

²*Department of Political Science, University of California, San Diego*

³*School of Cybersecurity and Privacy, Georgia Institute of Technology*

⁴*Department of Political Science, Vanderbilt University*

March 10, 2023

Abstract

Defense policy makers have become increasingly concerned about conflict in the “gray zone” between peace and war. Such conflicts are often interpreted as cases of deterrence failures, as new technologies or tactics—from cyber operations to “little green men”—seem to increase the effectiveness of low-intensity aggression. However, gray zone conflict could also be a case of deterrence success, where challengers adopt a constrained form of aggression in response to a credible escalation threat. We develop a model that formalizes both scenarios and identifies distinct empirical patterns across the two cases. We use the model’s findings to empirically analyze Russian gray zone activity since the 1990s, finding that Russian activity appears, in part, to be restrained by NATO’s deterrent threat. Our model also shows that developing gray zone conflict capabilities can lead to more peace but could also backfire and provoke a challenger to escalate to war.

*Corresponding author: andres.gannon@gmail.com. The authors wish to thank the members of the Center for Peace and Security Studies (cPASS), particularly Rex Douglass, as well as Nadiya Kostyuk, Kendrick Kuo, Chad Levinson, Sara Plana, Sanne Verschuren, John Warden, and Christopher Whyte for thoughtful feedback. Tom Brailey, Cole Reynolds, Benjamin Smalley, and Erin Werner provided excellent research assistance. Earlier drafts of this paper were presented at the 2018 STRATCOM Deterrence Symposium and the 2018 American Political Science Association conference. This research was sponsored by Office of Naval Research Grant N00014-14-1-0071 and the Department of Defense Minerva Research Initiative. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. government. Data and code for the empirical analysis can be found at <https://github.com/CenterForPeaceAndSecurityStudies/GrayZone>.

1 Introduction

After the downfall of Ukrainian President Viktor Yanukovich in February 2014, the Crimean Peninsula was invaded by “little green men,” Russian soldiers whose uniforms lacked insignia or other identifying information. The Kremlin formally annexed Crimea shortly thereafter. Protracted ground skirmishes, cyber campaigns, and “active measures” continued to plague Ukraine for years. Russia’s intervention has emerged as a paradigmatic example of a technologically novel and politically efficient form of “hybrid warfare,” designed to challenge the status quo without triggering broader conflict (Marten 2015; Lanoszka 2016). Similar tactics and imagery have emerged elsewhere, like Chinese “little blue men” eroding “red lines” in maritime East Asia (Green et al. 2017). According to former British Defense Secretary Michael Fallon (2017), “That is not a Cold War. It is a grey war. Permanently teetering on the edge of outright hostility. Persistently hovering around the threshold of what we would normally consider acts of war.”

There is increasing concern that conventional conceptions of deterrence are inadequate to address burgeoning threats in the gray zone (Holmes and Yoshihara 2017; Matisek 2017; Hicks and Friend 2019; Pettyjohn and Wasser 2019). Deterrence is typically believed to have failed if a challenger disrupts the status quo or resorts to military violence. From this perspective, traditionally capable defenders appear ill-equipped to respond to revisionism in the gray zone (Jackson 2017). As General Dunford (2016), Chairman of the United States Joint Chiefs of Staff, commented, “Our traditional approach is either we’re at peace or at conflict. And I think that’s insufficient to deal with the actors that actually seek to advance their interests while avoiding our strengths.”

In contrast to the concerns raised above, we suggest that gray zone conflict is neither novel nor inherently a deterrence failure. Rather than repudiating deterrence, gray zone challenges could actually reflect a certain respect for existing frameworks, both to avoid major war and preserve interests in common with status quo powers. Deterrence not only determines *whether* a challenge emerges but also *how*. Restricting other nations’ behaviors could certainly prove valuable, even in the absence of peace and a full retention of the status quo. An enemy that engages “with one hand tied behind its back” to avoid triggering a larger contest is not fighting effectively. Even if the challenger resorts to force and the defender does not escalate, fear of subsequent intervention

by the defender could cause the challenger to adopt a more furtive military strategy.

The research on deterrence is vast and heterogeneous (Huth 1999; Freedman 2004; Danilovic and Clare 2010; Quackenbush 2011). Yet there is relatively little work on choices about the *means* of deterrence compared to the literature on political *ends* (Carcelli and Gartzke 2017). We contribute by developing a formal model where a challenger and defender can select from variable intensities of limited conflict or resolve the crisis with a decisive war. We find that actors will behave differently depending on if the actor is constrained by their internal cost-benefit balancing (i.e. driven by costs, efficacy, innovation, etc.), or if the actor is constrained by an external deterrent threat. We then derive a series of Observations that connect empirically observable behavior to the challenger’s constraints and motivations for gray zone conflict (i.e. whether a challenger is constrained internally or externally; Section 6). We find evidence that the external deterrent threat of NATO intervention seemed to shape recent Russian uses of gray zone conflict (Section 7).

Finally, we discuss an additional result from our formal model: that a defender becoming better at gray zone conflict can incentivize a challenger to risk war. This result arises because as the defender gets better at fighting within gray zone conflict, gray zone conflict becomes less beneficial to the challenger. For a challenger who would otherwise conduct gray zone conflict, this change in the defender’s gray zone capabilities can result in the challenger giving up on gray zone conflict altogether and accepting the status quo, or (more surprisingly) can result in the challenger acting more aggressively and risking war. While the case is still evolving at the time of writing, we discuss how our result could offer an explanation for the February 2022 Russian invasion of Ukraine (Section 8).

2 Between Peace and War

Despite the analytical convenience afforded by conceptualizing war and peace as discrete outcomes, there is nothing new about conflict that falls between these two extremes. The literature on conflict is replete with both examples and frameworks, including limited war (Kissinger 1955; Osgood 1969), salami slicing (Schelling 1966), low-intensity conflict (Turbiville 2002), hybrid wars (Lanoszka 2016), frozen conflict (Driscoll and Maliniak 2016), covert operations (Carson 2018; O’Rourke 2018),

cyberwar (Lindsay and Gartzke 2018), and hassling (Schram 2021).

Early Cold War writings on “weakening the enemy with pricks instead of blows” emphasized that states may choose to pursue limited political objectives in the shadow of nuclear escalation (Hart 1967, 186). The Korean War seemed a then-underappreciated type of war fought to achieve political ends short of total victory with military means short of a total commitment (Osgood 1969). Contemporary treatments understood limited war as a conflict between actors that had the capacity to increase battlefield commitments but did not want to do so, creating a third option between major war and acquiescence (Brodie 1957; Kissinger 1957). Strategists introduced the “stability-instability paradox” to describe how disincentives for nuclear war, or even major conventional war, encourage conflict at lower levels of intensity, or in peripheral theaters (Snyder 1965; Jervis 1984). Similarly, Powell (2015, 598) notes that “how much power the challenger brings to bear limits how much risk the defender can generate.” As George and Smoke (1989, 173) explain, adversaries can “design around” deterrence by discovering new options that offer “an opportunity for gain while minimizing the risk of an unwanted response”.

The renewal of interest in low-intensity conflict between more capable competitors represents a return to these themes. A common thread in definitions of gray zone conflict is that it involves “a carefully planned campaign operating in the space between traditional diplomacy and overt military aggression” employed by challengers with grand geopolitical ambitions and potent capabilities (Mazarr 2015).¹ A number of practitioners and journalists highlight the worrying expansion of technologies by which low-intensity conflict can be practiced (Olson 2015). Scholars similarly note the technological novelty of this phenomenon as a form of aggression against which the US is unprepared (Hoffman 2007; Thornton 2015; Brands 2016; Jackson 2016; Wirtz 2017; Hughes 2020). This pessimism concerning adversaries’ ability to overthrow the existing military balance with “innovative doctrines or cunning strategies” (Goldstein 2017, 906) has even led some to advocate revamping deterrence to focus on gray zone threats (Tor 2015; Matisek 2017). The policy prescription emerging from this conventional wisdom is that the United States—and other targets of gray zone aggression—must use more of the tools available at its disposal to re-assert deterrence and

¹There is much ambiguity concerning how practitioners define and interpret gray zone conflict. See Bragg (2017) and Janičatová and Mlejnková (2021).

counter innovations in this unconventional battle space (Hicks and Friend 2019; McCarthy, Moyer and Venable 2019).

We wish to push back on this growing consensus that an expanded repertoire of potent tools for conflict short of war means gray zone conflict represents the new modal form of conflict. Indeed, militaries have been innovating novel technologies in *all* domains of warfare, at *all* levels of intensity (Gannon 2022). Yet in most conflicts, many of the sharpest arrows remain in the quiver. Therefore, the choice to use select innovations but not others represents a restriction, rather than an expansion, of the total means available. This shift in perspective, which views gray zone conflict as a relative reduction rather than an absolute expansion of options, calls into question claims about its effectiveness and even its novelty.

In the last decade of the Cold War, US Secretary of State George Shultz (1986, 204) offered a note of cautious optimism in this regard:

The ironic fact is, these new and elusive challenges have proliferated, in part, because of our success in deterring nuclear and conventional war. Our adversaries know they cannot prevail against us in either type of war. So they have done the logical thing: they have turned to other methods. Low-intensity warfare is their answer to our conventional and nuclear strength a flanking maneuver, in military terms.

Below we develop Shultz’s insight to explore whether modern gray zone conflict should be viewed as similarly reassuring or newly alarming.

3 Theoretical Intuition

The Euromaidan Revolution (2013-2014) presented Russia with a new political reality: Kiev was abruptly realigning with the West. In response, Russia quickly set about attempting to revise this realignment. From 2014 to late 2021, Russia’s actions were finite and often indirect, far below the threshold of Russian military capabilities. We refer to behaviors like these as limited “challenges,” implicitly to the status quo.² In response to Russian challenges, NATO quickly increased its presence

²The notion of “status quo” is inherently contextual. We treat the status quo as whatever political conditions prevail at a given point in time.

in the Black Sea, reinforced its support for capacity-building in Ukraine, and stepped up its presence and cooperation in other countries in Eastern Europe. This helped Ukraine counter the Russian-sponsored separatist movement in eastern Ukraine, while Russian cyber campaigns failed to influence the political alignment of Kyiv. In this way, Ukraine became engaged in a gray zone conflict.

We model these dynamics in terms of a challenger and a defender in a crisis.³ First, the challenger decides whether to accept the status quo and do nothing, or to conduct an opportunistic challenge selected from a continuum. If the challenger conducts a challenge, then the defender can respond by accepting, by escalating to a decisive war, or by countering the challenge with limited conflict. If the defender engages in limited conflict, the game proceeds and the defender makes an ultimatum offer to the challenger, who can accept the offer or reject it and go to war. When the challenger engages in a challenge, the defender responds with a limited response, and the challenger does not escalate to war, we will say that the states are engaged in a gray zone conflict. Thus, *gray zone conflict occurs when a militarily capable challenger intentionally limits the intensity and capacity with which they conduct military operations, and the defender engages but chooses not to escalate to a decisive war*. Our definition means that gray zone conflict must be preferred by both sides in a contest. Both actors have the capacity to escalate to a larger war but both also prefer not to, meaning that gray zone conflict is an equilibrium (Carson 2016).

Our model highlights the trade-offs that actors confront when choosing varying degrees of force. An actor may forgo the most decisive means when some other more appealing alternative exists. While war can accomplish an aggressor's goals, it could also be unnecessary and inefficient if partial victories or *faits accomplis* can be achieved at lower cost (Tarar 2016; Altman 2018). By considering these trade-offs, our model offers insight into two central questions. First, why do states engage in gray zone conflict and, once they do, what explains variation in the intensity with which they pursue it? Second, when a defender faces the prospect of challenges, can a robust ability to resist these challenges prove counterproductive for the defender?

Challengers that attempt to alter the status quo must choose a level of aggression. A more intensive challenge can yield greater political gains, but could introduce costs. The selected degree of challenge

³Many gray zone challenges involve either state or non-state proxies. For parsimony, we consider a target's allies as part of the target's capabilities.

faces two possible constraints: an *external deterrent constraint*, and an *internal efficiency constraint*. If the *external deterrent constraint* binds, the challenger is resolved, but will limit their level of force to avoid a greater risk of war. Here the challenger is most reactive to the defender’s willingness to go to war. Alternatively, if the *internal efficiency constraint* binds, the challenger restricts their level of force based on their own internal cost-benefit calculations, weighing the rewards and risks of gray zone conflict. Here the challenger is most sensitive to its own valuation of the stakes and the costs of its challenge.

The theoretical distinction described above is empirically consequential. The intensity of gray zone conflict is shaped by one of the two distinct constraints described above. Our model establishes that, depending on what constraint binds, we may observe distinct comparative statics. When the challenger’s internal efficiency constraint binds, the challenger freely chooses the scope and intensity of conflict that it believes is most efficient for accomplishing its objectives without concerns of provoking an escalation. When the challenger’s external deterrent constraint binds, by contrast, the challenger must scale back from its optimal low-level challenge in order to avoid triggering a larger contest. Our model and results allow us to connect observable relationships in the data to what constrains and motivates a challenger’s gray zone conflict decision. A key implication of these results and our empirical analysis is that we offer suggestive evidence that Russia may be constrained by NATO’s external deterrent threat in Russia’s conduct of gray zone conflict.⁴

When is a robust ability to counter challenges beneficial for the defender? When a defender is effective at fighting gray zone conflict, gray zone conflict becomes less efficient for the challenger. This can encourage the challenger to pursue options outside of gray zone conflict, producing mixed results for the defender state. On one hand, becoming effective at gray zone conflict could be productive for the defender if, upon the challenger’s low-level activity becoming ineffective, the challenger abandons the use of force altogether and accepts the status quo. On the other hand, the defender becoming better at gray zone conflict could be counterproductive if, upon the challenger’s low-level activity becoming ineffective, the challenger instead escalates to war. We formalize and elaborate on this point below, and we discuss it in the context of the February 2022 Russia invasion of Ukraine (see Section 8).

⁴This result is subject to standard endogeneity concerns in the multiple regression setting.

4 Model Setup

Two actors, a challenger (C) and a defender (D), are in a crisis over a divisible asset with a normalized value of 1. C wants to select some challenge (g_C) to improve its future political outcomes. In response, D will accept the challenge, go to war before the challenge is realized, or select a gray zone response (g_D). When D does not declare war, g_C and g_D affect future bargaining between C and D by influencing future wartime payoffs. At the end of the game, D offers C some political concession, and C can accept or go to war over the issue. D's decision making is influenced by D's private costs from war. In this setup, C does not know how great a challenge it can make before D might go to war. The game is as follows.

1. Nature sets $\kappa_D = \underline{\kappa}_D$ with probability ω and sets $\kappa_D = \bar{\kappa}_D$ with probability $1 - \omega$, with $0 \leq \underline{\kappa}_D < \bar{\kappa}_D$. D knows nature's selection of κ_D , but C does not.
2. C selects challenge level $g_C \geq 0$.
3. D can either go to war by setting $w_D = 1$, or not go to war by setting $w_D = 0$ and selecting some gray zone response $g_D \geq 0$. Going to war terminates the game and produces wartime payoffs $U_C = \theta P(0, 0) - \kappa_C - \beta_C g_C^2$ and $U_D = 1 - P(0, 0) - \kappa_D$ for C and D (respectively). When D does not go to war, the game moves to the next stage.
4. D offers C $x \in [0, 1]$.
5. C can declare "war" by setting $w_C = 1$ or can "accept" the offer by setting $w_C = 0$. When C sets $w_C = 1$, C receives payoff $U_C = \theta P(g_C, g_D) - \kappa_C - \beta_C g_C^2$, and D receives $U_D = 1 - P(g_C, g_D) - \kappa_D - \beta_D g_D^2$ (respectively), which is their "updated" wartime payoffs. When C sets $w_C = 0$, C receives payoff $U_C = \theta x - \beta_C g_C^2$ and D receives $U_D = 1 - x - \beta_D g_D^2$.

To offer some intuition, Stages 1-3 describe C's challenge and D's initial response, and Stages 4-5 describe future crisis bargaining after C's challenge and D's gray zone response has been realized. To offer empirical grounding, Chinese behavior in the South China Sea (circa 2022) is a challenge ($g_C > 0$), and neighboring countries responding to these challenges (whether through indirect measures like patrols, or through direct confrontation) is the gray zone response ($g_D > 0$). This

gray zone activity (when $g_C > 0$ and $g_D > 0$) will influence whatever eventual political settlement, stalemate, or later war that emerges between China and its neighbors (Stages 4-5). Payoffs are summarized in Table 1.

Gameplay	C's utility	D's utility
<i>D initiates war at stage 3</i> ($w_D = 1$ before g_C and g_D are realized)	$\theta P(0, 0) - \kappa_C - \beta_C g_C^2$	$1 - P(0, 0) - \kappa_D$
<i>C initiates war at stage 5</i> ($w_C = 1$, after g_C and g_D are realized)	$\theta P(g_C, g_D) - \kappa_C - \beta_C g_C^2$	$1 - P(g_C, g_D) - \kappa_D - \beta_D g_D^2$
<i>C accepts at stage 5</i> (after g_C and g_D are realized)	$\theta x - \beta_C g_C^2$	$1 - x - \beta_D g_D^2$

Table 1: Summarized payoffs for actors.

The function $P(\cdot, \cdot)$ is the likelihood that C wins in a war. We assume functional form $P(g_C, g_D) = \max \{ \min \{ 1, \rho + g_C - g_D \}, \rho \}$ with $\rho \in [0, 1]$, implying that P is weakly increasing in g_C , weakly decreasing in g_D , and that P falls between ρ and 1 inclusive. The constant ρ is C's likelihood of winning a war before any gray zone activity comes to fruition (when $g_C = 0$ and $g_D = 0$). Note that this is the likelihood that C wins in a war when D initiates war in Stage 3. The assumption that $P(t, h)$ must be weakly greater than ρ implies that gray zone responses by D can undercut challenges by C but cannot make C a declining power, at most returning C to a baseline war victory likelihood of ρ .⁵ Challenges and gray zone responses are costly, and these costs are captured in the $\beta_C g_C^2$ and $\beta_D g_D^2$ terms (respectively), where $\beta_D > 0$ and $\beta_C \geq 0$.

If D initiates war in Stage 3, then C and D fight over the asset. C's likelihood of winning in war is $P(0, 0) = \rho$, and $\kappa_C > 0$ and $\kappa_D > 0$ are C's and D's costs from war, respectively. Regarding costs $\kappa_D \in \{\underline{\kappa}_D, \bar{\kappa}_D\}$, we adopt a simple functional form, assuming that $\underline{\kappa}_D = \kappa$, and $\bar{\kappa}_D = \kappa + \psi$, with $\psi > 0$. This will be relevant below where we will manipulate D's willingness to go to war across types (i.e. when we manipulate D's deterrent threat, or κ).⁶ If C initiates war in Stage 5,

⁵While low-level conflict could potentially turn an actor into a declining power, this falls outside of the gray zone conflict framework (and larger deterrence framework) that we adopt here. That said, the key results—for example, that improving gray zone capabilities could lead to worse outcomes for D (Observation 7)—would survive if g_D could make C decline, but under different conditions.

⁶The functional form for κ_D adopted here assists in interpreting comparative statics. While D's true type (whether D is type $\underline{\kappa}_D$ or $\bar{\kappa}_D$) is unknown to C, C does know the value of κ .

the fighting happens after the gray zone activity has occurred, which means that P now depends on g_C and g_D . If C accepts in Stage 5, this represents a settlement after the gray zone activity has occurred, and value x denotes the offer D makes to C.

Readers might take issue with the notion that challenges shift the future war-fighting capabilities of the actors or with the structure of future bargaining between the actors. The model here is isomorphic to an alternate framing where C's and D's gray zone activity (g_C and g_D) shape some future payoff between the actors that is outside of either a negotiated peace or war (similar to models in Tarar (2016) and Gurantz and Hirsch (2017)).⁷ Additionally, similar results have been generated in game-form free environments (Kenkel and Schram 2021).

Readers might also take issue with us assuming that D's costs of gray zone conflict do not influence D's costs of war.⁸ We assume this as a technical simplification here, but include a model in Appendix Section 2 where the costs of war and gray zone conflict are correlated. This model with correlated costs produces substantively similar results, with some caveats and technical distinctions. Ultimately, we believe the model with correlated costs in the Appendix best models reality, but we also believe that the model presented in the main text, where costs are not related, is not entirely off base. After all, there are many differences to how states conduct gray zone conflict and war, and these different types of fighting can generate very different costs (Kenkel and Schram 2021).

As some evidence of this, we can look at the conflict in Ukraine before and after the 2022 Russian escalation. Before the invasion, Russia was engaged in an externally sponsored insurgency that sought to limited overt direct involvement. Upon invading Ukraine in February 2022, support for indigenous military forces has not been a central priority; unmarked Russian Spetsnaz teams no longer dominate the urban battlefields, and obscuring attribution is no longer emphasized. Supporting a third-party militant group has distinct limitations and costs relative to using one's own soldiers. On many other dimensions, the war looks different for Russia. Recently, Russia is facing manpower shortages and domestic discontent over conscription, and recently implemented a series of domestic financial interventions to prop-up struggling local markets. These are problems that

⁷As one way to do this, the game could end in Stage 3 where, should gray zone conflict occur (rather than war or peace), C and D receive payoff functions $X_C(g_C, g_D)$ and $X_D(g_C, g_D)$. This could allow more flexibility to payoffs, but would not include bargaining between actors.

⁸Tarar (2016) models the costs from *fait accompli* in a similar way.

did not exist during Russia’s gray zone operations, meaning Russia’s ability to weather them is orthogonal to an ability to fight an effective proxy war. We discuss this more (and analyze the model with correlated gray zone and war costs) in Appendix Section 2.

The model integrates features from the formal literature on endogenous power shifts and deterrence (Fearon 1997; Tarar 2016; Debs and Monteiro 2014; Gurantz and Hirsch 2017; Baliga, Mesquita and Wolitzky 2020) while pushing the decision-theoretic strands of classical deterrence theory towards embracing a broader repertoire of possible actions (Zagare 1996; Huth 1999). It extends existing research by considering two competing actors, each with a continuum of conflict options outside of the standard crisis-bargaining moves of declaring war or accepting peace. Far from just a technical flourish, that each actor can select some low-level of conflict is central to how we define gray zone conflict and to some of our more counter-intuitive results (like Observation 7). The model is similar to those where actors have flexible responses rather than just war or peace (Zagare and Kilgour 1998; Werner 2000; Schultz 2010; Slantchev 2011; Powell 2015; Tarar 2016; McCormack and Pascoe 2017; Coe 2018; Spaniel 2019; Schram 2021). This model is most similar to that in Schram (2021) but is still unique in several key respects; (1) this model assumes a different kind of information asymmetry (the costs of war are uncertain); (2) C faces costs to conducting challenges; and (3), we consider new parameters like C’s resolve. For these reasons, the results and comparative statics derived here are new.

4.1 Equilibrium Concepts and Assumptions

We limit attention to pure strategy perfect Bayesian Nash equilibria. As we discuss in the Appendix, considering mixed strategies does not change the substance of the results. We let x^* , g_C^* , g_D^* , w_C^* , and w_D^* denote equilibrium actions.

The scope of the analysis concerns scenarios where D is willing to go to war over significant challenges, where it is feasible for C to select a challenge that C knows will end in gray zone conflict, and where there could be positive returns for C to risk war. Formally, these assumptions are $1 - \rho - \bar{\kappa}_D > 1 - \frac{\kappa_C}{\theta}$, $-\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta} \geq 0$, and $(1 - \omega) \left(-\frac{1}{4\beta_D} + \bar{\kappa}_D \right) + \frac{1}{4\beta_D} - \underline{\kappa}_D \geq 0$ (respectively). The first assumption means that there is some g_C value that will result in D responding with war.⁹

⁹Formally, this assumption means that D’s payoff from going to war in the third round ($1 - \rho - \bar{\kappa}_D$) is greater

Essentially, the first assumption rules out cases where D does not care much for the asset, and where C is able to do exactly what they want without any response. This could describe a case where C undertakes some domestic policy that D dislikes, but where D is not willing to impose economic or military punishments when C does so. The second assumption means that gray zone conflict is feasible, and that C can conduct challenges that will end in gray zone conflict with certainty.¹⁰ Essentially, the second assumption rules out empirical cases where D cares too much for the asset (relative to D's costs of war), meaning gray zone conflict would never actually occur on the equilibrium path. The third assumption implies there could be positive returns for C to risk war; in other words, the $1 - \omega$ chance of getting a greater gray zone conflict payoff is worthwhile.¹¹ Because we are considering empirical cases where war could occur, we view this as reasonable.

Finally, for the statement of equilibrium strategies below, we also assume that $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$.¹² This assumption means that C is capable enough at conducting a limited response that doing so would be productive for C. Without this assumption, whenever $\frac{\theta}{2\beta_C} \leq \frac{1}{2\beta_D}$, C will never challenge. This assumption makes little substantive difference, and all Observations below are proved without it.

5 Equilibrium

5.1 Equilibrium Intuition

Working backwards, in the fourth and fifth stages, D does strictly better avoiding war, and D will make an offer that makes C indifferent between accepting and fighting. When these stages are reached, in equilibrium, $x^* = P(g_C^*, g_D^*) - \frac{\kappa_C}{\theta}$ and $w_C^* = 0$. In the third stage, D will see the selected challenge g_C^* and will respond with acceptance, a gray zone response, or war. D's response

than D's payoff from C selecting some very large g_C , D not selecting any gray zone response ($g_D = 0$) or war, and in the fourth stage setting $x = 1 - \frac{\kappa_C}{\theta}$.

¹⁰Formally, this assumption implies that if C had very low costs to gray zone conflict, C could select $g_C^* = \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}$, D would optimally respond with $g_D^* = \frac{1}{2\beta_D}$, and this would result in C doing better in bargaining in rounds 4-5 relative to if C selected $g_C^* = 0$.

¹¹Note that we are not claiming that C always prefers war or that war is not still inefficient. As we will show later, C's payoff from going to war sometimes when the deterrent constraint binds is $(1 - \omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) + \omega(\theta\rho - \kappa_C) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ and C's payoff from always selecting into gray zone conflict is $\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$. If the condition in the text did not hold, then $(1 - \omega)\theta\left(-\frac{1}{4\beta_D} + \bar{\kappa}_D\right) < \theta\left(-\frac{1}{4\beta_D} + \bar{\kappa}_D\right)$, implying that C would always prefer gray zone conflict with certainty to ever risking war; when this does not hold, then C may still not prefer risking war sometimes (this inequality says nothing of C's costs to the challenges).

¹²When $\beta_C = 0$ this constraint becomes irrelevant.

is a function of the scope of C's challenge (g_C^*), D's costs of war ($\underline{\kappa}_D$ or $\bar{\kappa}_D$), and the costs and benefits of conducting a gray zone response (β_D and the eventual x^*). D will go to war ($w_D^* = 1$) if the challenge is large, D's costs of war are low, D's utility from gray zone conflict is low, or some combination of these. Otherwise, D will select some limited response, which takes the form $g_D^* = \min\{g_C, \frac{1}{2\beta_D}\}$. Note that when $g_C = 0$, D will avoid conflict altogether and set $g_D^* = 0$.

C's action in the second stage determines how the rest of the game plays out. The game could end peacefully. For example, if C's costs of conducting challenges are very high, C will accept the status quo ($g_C^* = 0$). Doing so will result in D not engaging in a gray zone response (as there is no challenge to undercut), and the game will end with no costly conflict.

The game could also end with gray zone conflict. Suppose C wants to select the most aggressive challenge possible without provoking D to go to war. To avoid war altogether, C must make sure that the type of D that is the best at war, type $\underline{\kappa}_D$, is unwilling to escalate. So long that C selects a g_C where $g_C \leq \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$, then type $\underline{\kappa}_D$ weakly prefers not going to war. Of course, C may not always want to select the most aggressive challenge, as C faces greater costs for doing so. C also faces an internal optimization tradeoff based on C's resolve and C's costs for conducting challenges—essentially where marginal returns to challenges are equal to marginal costs. Here C optimizes their payoffs by setting $g_C = \frac{\theta}{2\beta_C}$, so long that this value overtakes D's anticipated limited response.¹³ These values of g_C represent constraints. Whenever $\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta} < \frac{\theta}{2\beta_C}$ and C does not wish to go to war, C would like to select a more aggressive challenge based on C's internal cost-benefit calculations, but C is constrained by the threat of provoking D to war. Here, the *external deterrent threat* constraint binds, and C sets $g_C = g_C^{GZ,EDT} = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$. The external deterrent threat constraint binds when C has high resolve and low costs to conducting challenges, and when D is less effective at war and gray zone conflict. Alternatively, whenever $\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta} \geq \frac{\theta}{2\beta_C}$, C can select their best optimal challenge (i.e. C is constrained by their own costs) without worrying about provoking D to go to war. Here, the *internal efficiency* constraint binds, and C sets $g_C = g_C^{GZ,IE} = \frac{\theta}{2\beta_C}$. Note that when $\beta_C = 0$, C has no costs to conduct challenges and the internal efficiency constraint never binds.

¹³Because D optimally sets $g_D = \frac{1}{2\beta_D}$, so long that $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ C's challenge shifts P in C's favor. Also note, if setting $g_C = \frac{\theta}{2\beta_C}$ ever results in war, a different optimization must be considered. Finally, whenever $\beta_C = 0$, the internal efficiency constraint never binds.

Finally, C may want to risk war. This can arise through the standard risk-return tradeoff that actors face in crisis bargaining models (Fearon 1995; Ramsay 2017), only here C risks war to undertake more aggressive challenges. By setting $g_C \leq g_C^{GZ,EDT}$, C avoids war altogether. But, C may want to select a more aggressive challenge ($g_C > g_C^{GZ,EDT}$) in order to conduct more aggressive challenges against privately-bad-at-war types of D (type $\bar{\kappa}_D$), an action that would risk war if D turns out to be a privately-good-at-war type (type $\underline{\kappa}_D$). While C does not know what type D is at this point in the game, the expected payoffs from selecting some more aggressive challenge ($g_C > g_C^{GZ,EDT}$) that results in war with probability ω may be worthwhile, especially when ω is small. When C risks war, then C faces an analogous set of constraints as C did within gray zone conflict. To keep both types of D from going to war, C must limit their challenge. This creates another external deterrent threat constraint (here for type $\bar{\kappa}_D$ D's), where $g_C^{W,EDT} \leq \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$. C similarly wants to optimize their returns on challenges given their costs. This creates another internal efficiency constraint, where $g_C^{W,IE} = \frac{(1-\omega)\theta}{2\beta_C}$. Together, within the parameters where C optimally risks war, C's external deterrent threat constraint binds when $g_C^{W,EDT} < g_C^{W,IE}$, resulting in C setting $g_C = g_C^{W,EDT}$. C's internal efficiency binds when $g_C^{W,EDT} \geq g_C^{W,IE}$, resulting in C setting $g_C = g_C^{W,IE}$.

C's selection across these five potential challenges (i.e. selecting $g_C^* \in \{0, g_C^{GZ,IE}, g_C^{GZ,EDT}, g_C^{W,IE}, g_C^{W,EDT}\}$) is determined by what constraints bind and what utilities the status quo, gray zone conflict, or war give C.¹⁴ Figures 1 and 2 visualize how, under select parameters, each of these possible g_C values can be an equilibrium.

First consider the top panel within Figure 1 (with the title “ $\beta_C = 3$: C Optimally Selects GZC, Deterrent Threat Binds”). For a fixed set of parameters,¹⁵ this figure displays C selecting some g_C on the x-axis, and C's final expected payoff given D's and C's best responses following C's choice of g_C on the y-axis. C's equilibrium selection, g_C^* , is denoted by the asterisks.

In the top panel, when $g_C = 0$, C selects into the status quo, and D neither wants to go to war ($w_D^* = 0$) nor wants to implement any costly response ($g_D^* = 0$). Later in the game, D makes an offer that gives C their wartime utility, and C will accept.

¹⁴C's utilities are listed in the Appendix.

¹⁵The parameters for all figures are listed in the Appendix.

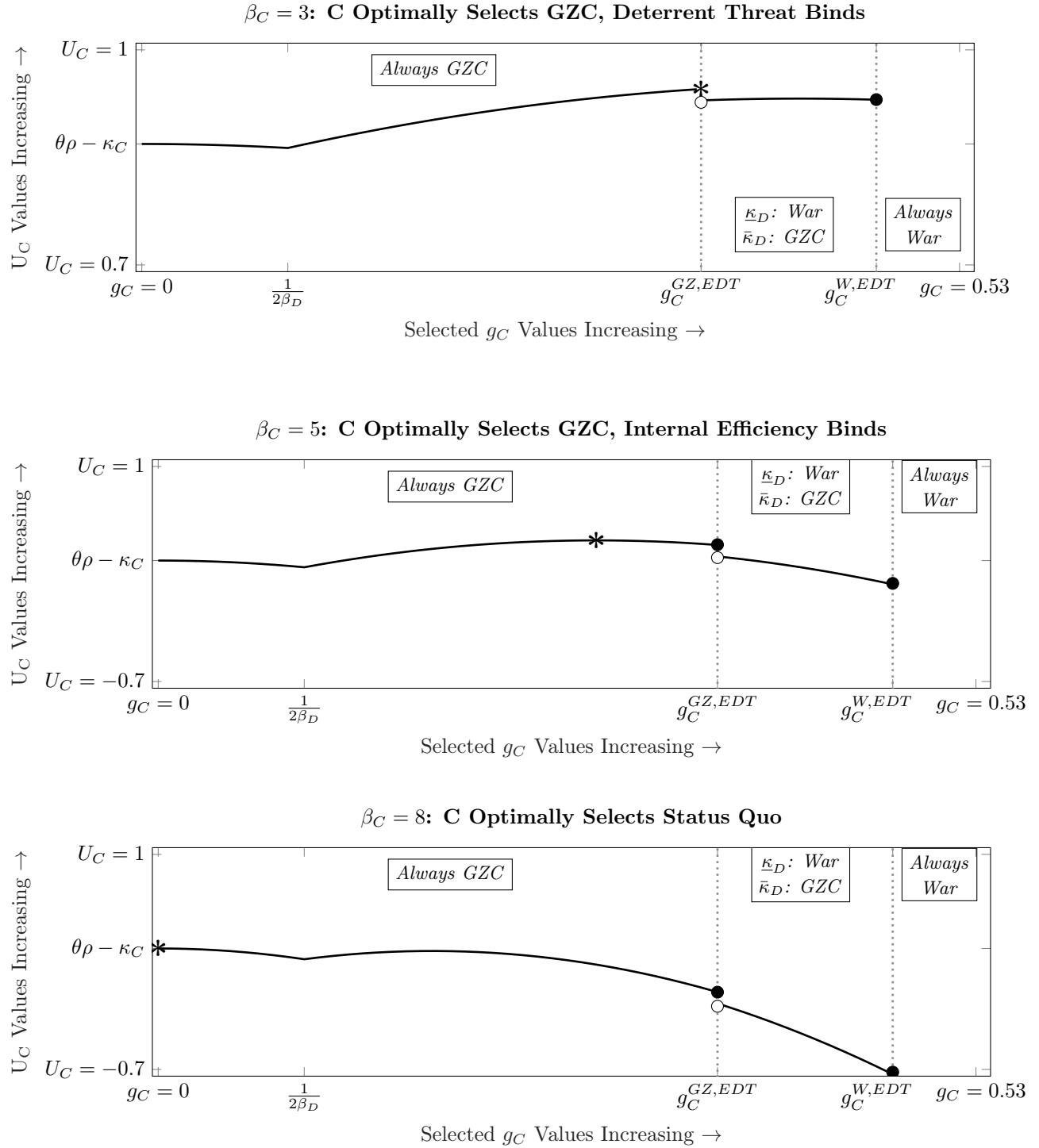


Figure 1: C's Utility Across selected g_C 's (Status Quo & Gray Zone Conflict Outcomes).

C's optimal g_C is marked by the asterisks. The regions denote how D responds to a selected g_C —whether D will always engage with a gray zone response, whether D will always go to war, or whether different types of D respond differently. All sub-figures have the same underlying parameters (listed in the Appendix), except β_C which varies across sub-figures.

When $g_C \in (0, \frac{1}{2\beta_D}]$, C's utility is decreasing in g_C . For any g_C in this region, D will optimally respond with a gray zone response at level $g_D^* = \min \left\{ g_C, \frac{1}{2\beta_D} \right\}$, which effectively cancels out the benefits of C's challenge. And, because C is still paying costs for conducting the challenge, C does worse if C chooses to select greater challenges within this region.

When $g_C \in (\frac{1}{2\beta_D}, g_C^{GZ,EDT}]$, C's utility is increasing in g_C . Here, D's limited response ($g_D^* = \frac{1}{2\beta_D}$) is not enough to cancel out C's challenge, which means C's challenges start being productive. And, while C is facing costs to implementing more aggressive challenges, here β_C is not high enough to prevent C's utility from increasing (under these parameters, $g_C^{GZ,EDT} < g_C^{GZ,IE}$).

Moving to the right of $g_C = g_C^{GZ,EDT}$, different types of D begin to play the game differently, which creates a discontinuity. Type $\underline{\kappa}_D$ D will go to war for any g_C to the right of the threshold. And, type $\bar{\kappa}_D$ D (who is worse at war) is still willing to tolerate challenges slightly beyond $g_C^{GZ,EDT}$. Together, C's utility has a discontinuity because C is now going to war with probability ω , which is bad for C, and because C is not gaining much from small increases in g_C against type $\bar{\kappa}_D$. When $g_C \in (g_C^{GZ,EDT}, g_C^{W,EDT}]$, C's utility is increasing in g_C . So long that C's costs to g_C are low enough (as they are here: $g_C^{W,EDT} < g_C^{W,IE}$), C will do better selecting more aggressive challenges, as this will result in C doing better against type $\bar{\kappa}_D$.

Finally, there is another discontinuity at $g_C^{W,EDT}$, and for any $g_C > g_C^{W,EDT}$ C's utility is low and decreasing in g_C . Here both types of D will go to war, which is bad for C. Additionally, C incurs the costs of conducting the challenge while still ending up fighting a war with both types of D. Overall, these payoffs are (literally) off-the-chart bad, and for all parameters C prefers selecting into the status quo to any g_C in this space. Under this set of parameters, C attains the greatest utility by selecting $g_C^* = g_C^{GZ,EDT}$. Verbally, here C is willing to conduct the greatest challenge that will not push either type of D into war, resulting in an outcome of gray zone conflict.

Now consider the middle panel in Figure 1 (with the title “ $\beta_C = 5$: C Optimally Selects GZC, Internal Efficiency Binds”), where all parameters are the same except $\beta_C = 5$. Here C's costs of challenging have gone up, which alters the slopes of how g_C relates to C's payoffs. Now $g_C^{GZ,EDT} \geq g_C^{GZ,IE}$ holds, meaning C's internal efficiency constraint binds, and now C selects a challenge based on their own internal costs and benefits of challenging. Instead of C selecting the most aggressive

challenge they can that will not push D to declare war, C now does best stopping short of this value, instead selecting $g_C^* = g_C^{GZ,IE}$. Finally, in the bottom panel, we increase β_C again and fix $\beta_C = 8$. Here C's costs of challenging have increased so much that it is never productive for C to engage in any challenge, and C prefers selecting into the status quo.

Figure 2 considers a different set of parameters and illustrates other possible equilibria to the game. In the top panel (titled " $\beta_C = 1.8$: C Optimally Risks War, Deterrent Threat Binds"), C does best setting $g_C^* = g_C^{W,EDT}$. Even though by selecting $g_C^* = g_C^{W,EDT}$ C must fight if D is type $\underline{\kappa}_D$, which is bad for C, C is undertaking a more aggressive challenge against type $\bar{\kappa}_D$, which is good for C. Together, C's expected payoff from taking a more aggressive challenge and risking war is greater than C's expected payoff from never risking war ($g_C = g_C^{GZ,EDT}$) or selecting into the status quo ($g_C = 0$). Finally, in the bottom panel of Figure 2, C faces slightly greater costs to challenging (all parameters are the same except now $\beta_C = 2.5$), which results in inequality $g_C^{W,EDT} \geq g_C^{W,IE}$ holding; this means C will not select the greatest g_C possible that will keep type $\bar{\kappa}_D$ from declaring war, but will rather scale back their selected challenge due to their internal efficiency constraint.

5.2 Statement of Equilibrium

The equilibrium is as follows.

Stage 5: If $x \geq P(g_C, g_D) - \frac{\kappa_C}{\theta}$, then C sets $w_C = 0$. Otherwise, C sets $w_C = 1$. If C previously set $g_C = \min \left\{ \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}, \frac{(1-\omega)\theta}{2\beta_C} \right\}$, then C believes D is type $\bar{\kappa}_D$ with probability 1. Otherwise, C believes D is type $\bar{\kappa}_D$ ($\underline{\kappa}_D$) with probability $1 - \omega$ (ω).

Stage 4: D always sets $x = P(g_C, g_D) - \frac{\kappa_C}{\theta}$.

Stage 3: For $\kappa_D \in \{\underline{\kappa}_D, \bar{\kappa}_D\}$, if $1 - P(0, 0) - \kappa_D > 1 - P(g_C, \min \left\{ \frac{1}{2\beta_D}, g_C \right\}) + \frac{\kappa_C}{\theta} - \beta_D * \left(\min \left\{ \frac{1}{2\beta_D}, g_C \right\} \right)^2$, then D will set $w_D = 1$. Otherwise, D will set $w_D = 0$ and $g_D = \min \left\{ \frac{1}{2\beta_D}, g_C \right\}$.

Stage 2: C believes D is type $\bar{\kappa}_D$ with probability $1 - \omega$ and type $\underline{\kappa}_D$ with probability ω . C's selection of a challenge is as follows:

Sometimes War, Deterrent Threat Binds: C sets $g_C = g_C^{W,EDT}$ if $g_C^{W,IE} > g_C^{W,EDT}$, $\omega(\theta\rho - \kappa_C) +$

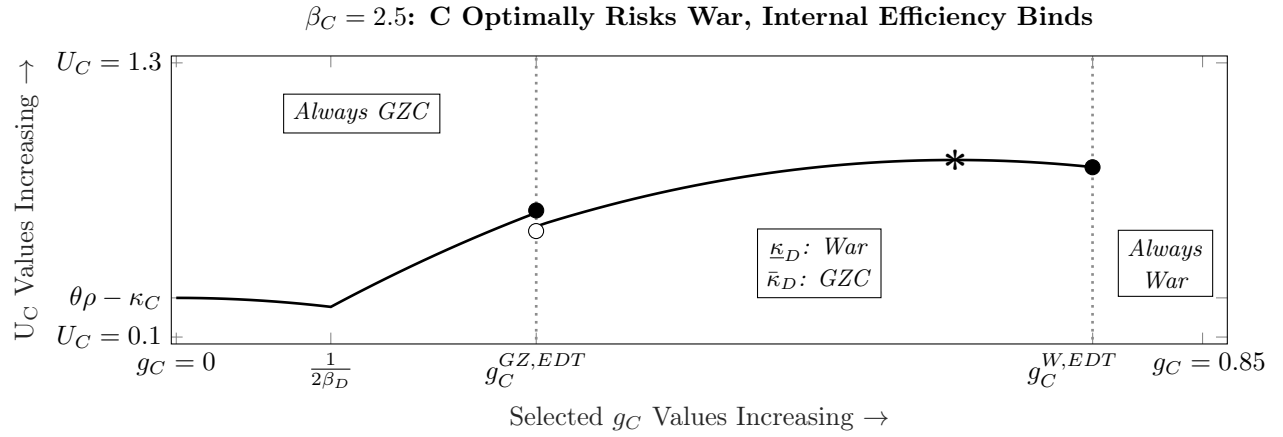
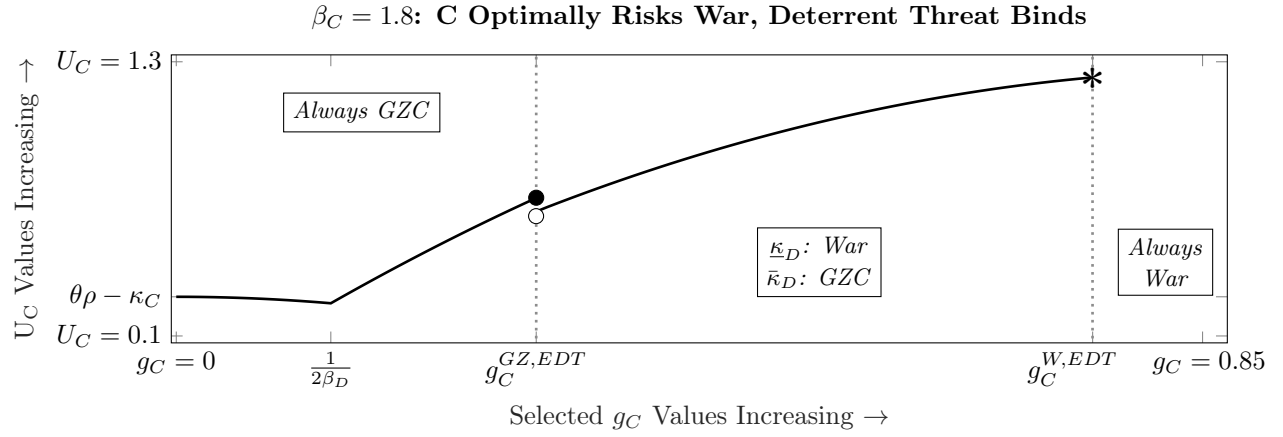


Figure 2: C's Utility Across selected g_C 's (War Outcomes).

C's optimal g_C is marked by the asterisks. The regions denote how D responds to a selected g_C —whether D will always engage with a gray zone response, whether D will always go to war, or whether different types of D respond differently. All sub-figures have the same underlying parameters (listed in the Appendix), except β_C which varies across sub-figures.

$$(1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(g_C^{W,EDT}\right)^2 > \theta\rho - \kappa_C, \text{ and } \omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(g_C^{W,EDT}\right)^2 > \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(g_C^{GZ,EDT}\right)^2.$$

Sometimes War, Internal Efficiency Binds: C sets $g_C = g_C^{W,IE}$ if $g_C^{W,IE} > g_C^{GZ,EDT}$, $g_C^{W,IE} \leq g_C^{W,EDT}$, $\omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(g_C^{W,IE}\right)^2 > \theta\rho - \kappa_C$, and $\omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(g_C^{W,IE}\right)^2 > \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(g_C^{GZ,EDT}\right)^2$.

Always Gray Zone, Deterrent Threat Binds: C sets $g_C = g_C^{GZ,EDT}$ if $\theta\rho - \kappa_C < \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(g_C^{GZ,EDT}\right)^2$ and any of the following holds: (a) $g_C^{GZ,IE} > g_C^{GZ,EDT}$ and $g_C^{W,IE} \leq g_C^{GZ,EDT}$; (b) $g_C^{W,IE} > g_C^{GZ,EDT}$, $g_C^{W,IE} \leq g_C^{W,EDT}$, and $\theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(g_C^{GZ,EDT}\right)^2 \geq \omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(g_C^{W,IE}\right)^2$; or (c) $g_C^{W,IE} > g_C^{W,EDT}$, and $\theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(g_C^{GZ,EDT}\right)^2 \geq \omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(g_C^{W,EDT}\right)^2$.

Always Gray Zone, Internal Efficiency Binds: C sets $g_C = g_C^{GZ,IE}$ if $g_C^{GZ,IE} \leq g_C^{GZ,EDT}$ and $\theta\rho - \kappa_C < \theta\rho + \frac{\theta^2}{4\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C$.

Status Quo: C sets $g_C = 0$ if any of the following holds: (a) $g_C^{GZ,IE} \leq g_C^{GZ,EDT}$ and $\theta\rho - \kappa_C \geq \theta\rho + \frac{\theta^2}{4\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C$; (b) $g_C^{GZ,IE} > g_C^{GZ,EDT}$, $g_C^{W,IE} \leq g_C^{GZ,EDT}$, and $\theta\rho - \kappa_C \geq \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(g_C^{GZ,EDT}\right)^2$; (c) $g_C^{W,IE} > g_C^{GZ,EDT}$, $g_C^{W,IE} \leq g_C^{W,EDT}$, $\theta\rho - \kappa_C \geq \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(g_C^{GZ,EDT}\right)^2$, and $\theta\rho - \kappa_C \geq \omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(g_C^{W,IE}\right)^2$; or (d) $g_C^{W,IE} > g_C^{W,EDT}$, $\theta\rho - \kappa_C \geq \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(g_C^{GZ,EDT}\right)^2$, and $\theta\rho - \kappa_C \geq \omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(g_C^{W,EDT}\right)^2$.

We discuss the equilibrium conditions in detail in the Appendix Section 1.3.

6 What Drives Variation in Gray Zone Activity?

Our model and its equilibrium are consistent with gray zone conflict being multifaceted, but add value by establishing two underlying, rational mechanisms for how challengers conduct gray zone conflict. First, the challenger may select their challenges based on the external deterrent threat, essentially pulling their punches in gray zone conflict to avoid a risk of war. Second, when the

external deterrent threat does not bind, the challenger selects their challenge based on their own internal efficiency, essentially choosing their optimal challenge unconstrained by the risk of further escalation. Those who claim that Russia (or any other state engaging in gray zone conflict) conducts whatever kind of gray zone conflict they want and operates outside the realm of standard deterrence dynamics are describing a challenger behaving based on their internal efficiency constraint. However, it could also be that Russia is limiting their challenges in order to avoid a greater escalation—which is more consistent with the external deterrent threat binding.

Here we establish how observable patterns can distinguish whether the challenger’s behavior is shaped by their internal efficiency or their external deterrent threat. This allows us to use the relationships uncovered by statistical analysis to suggest whether our challenger is behaving based on the logic of deterrence, or not. To be clear, these are not Observations describing a test of the model; rather, the statistical analysis in the next section is done to identify empirical patterns, which is then used as fodder for our theory (Hoover 2006; Clarke and Primo 2012). Put another way, in this Section, we provide a roadmap for how to use empirical findings that we can estimate to uncover the internal motivations for how challengers behave (i.e. what constraint binds).

6.1 On the Defender’s Willingness to go to War and Conflict Intensity

The observed relationship between the defender’s willingness to go to war, here approximated by κ , and the challenger’s selected challenges, g_C^* , depends on whether the challenger’s internal efficiency or the defender’s external deterrent threat shapes the challenger’s decision making.

When the external deterrent threat binds, the defender’s willingness to escalate to war creates an upper bound on the tolerated level of the selected challenge. If we observe a challenger decreasing their challenges as a defender becomes more willing to go to war, then we know that the challenger’s decisions are influenced by deterrence. For example, suppose under one set of parameters the challenger wants to select a challenge that will avoid war and where the external deterrent threat binds (formally $g_C^* = g_C^{GZ,EDT} = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$). If the defender becomes more willing to go to war and the challenger wishes to continue avoiding war, then the challenger becomes restricted in what challenges they can implement. This can play out in one of two ways. First, the challenger could continue selecting the challenge level that results in gray zone conflict ($g_C^* = g_C^{GZ,EDT}$), meaning,

as κ decreases, the challenger must decrease their challenge.¹⁶ Second, as the defender becomes more willing to go to war, the challenger could switch from selecting a challenge that results in gray zone conflict ($g_C^* = g_C^{GZ,EDT}$) to not challenging in the first place ($g_C^* = 0$). Together, when the challenger is responding to the external deterrent threat and is facing a defender that is more willing to fight, the challenger will pursue less aggressive challenges.

When the internal efficiency constraint binds, the challenger's costs and resolve determine the selected challenge. If we observe a challenger not altering their challenges as a defender becomes more willing to go to war, then we know the challenger is not responding to the defender's external threat from war, but rather is basing their decisions on their internal cost-benefit calculations. For example, suppose the challenger is selecting a challenge that will always result in gray zone conflict and where the internal efficiency constraint binds (formally $g_C^* = g_C^{GZ,IE} = \frac{\theta}{2\beta_C}$). If κ then decreases, as long as the internal efficiency constraint continues to bind, then there will be no changes in the selected g_C^* . Put another way, when the challenger is acting based on their internal efficiency and facing a defender that is more willing to fight, the challenger is non-reactive. Of course, it could also be that as κ decreases, the internal efficiency constraint no longer binds, and the external efficiency constraint starts to bind. When this is the case, the logic in the previous paragraph applies.

Consider Figure 3, which illustrates the dynamics described above. On the x-axis, left-to-right, the defender's deterrent threat from war is increasing (κ is decreasing). On the y-axis, we plot the challenger's selected equilibrium challenge g_C^* . For the greatest costs of war (the greatest κ), the challenger is unconstrained by the defender's external deterrent threat. This means that the challenger can freely select their challenge based on their internal cost-benefit balancing. Thus, the defender's costs of war do not play a role in the challenger's decision making here. Moving to the right (to the second region), the defender gets better at war, and the challenger cannot conduct the same level of challenge without provoking the defender. Here, the external deterrent threat starts to bind, and, to remain in gray zone conflict, the challenger must scale back their challenge, which results in g_C^* decreasing. Moving to the right further (to the third region), the defender is very willing to fight. Here the challenger is so constrained in what challenges they can conduct that will still result in gray zone conflict that the challenger no longer finds gray zone conflict to

¹⁶Put another way, $g_C^{GZ,EDT}(\kappa)$ is decreasing in its argument.

be productive. Under these very low values of κ , the challenger will select into the status quo ($g_C^* = 0$).

Altogether, Figure 3 illustrates a natural intuition. As the defender’s deterrent threat from war increases, the challenger cannot select as aggressive of challenges without provoking the defender to war. Decreasing κ results in a lower intensity of gray zone conflict until the challenger no longer finds gray zone conflict worthwhile.

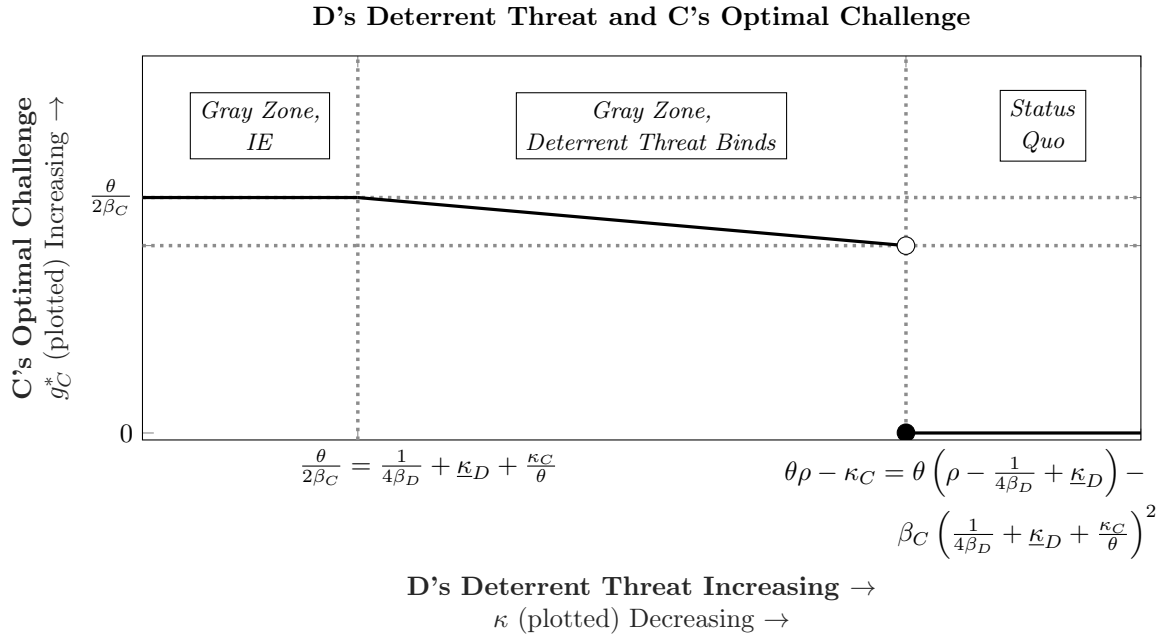


Figure 3: C’s Optimal Challenge as D’s Deterrent Threat Decreases.

C’s selected challenge under a range of κ ’s are plotted. “Gray Zone, IE” describes the equilibrium space where the outcome will always be gray zone conflict and C’s internal efficiency constraint binds. The first vertical boundary—where $\frac{\theta}{2\beta_C} = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ —defines where κ has increased to the cutpoint where D’s deterrent threat will start to constrain C’s gray zone behavior. The second horizontal boundary—where $\theta\rho - \kappa_C = \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2$ —is the cutpoint where D’s external constraint is so strong that C prefers accepting the status quo to engaging in gray zone conflict.

Two other dynamics are worth discussing. First, under the parameters in Figure 3, when κ decreases enough, the challenger switched into the status quo. Under different parameters, as κ decreases, the challenger may similarly feel constrained within gray zone conflict, but may switch to sometimes risking war.¹⁷ What distinguishes these cases is the challenger’s willingness to fight. When the

¹⁷Formally, when C conducts gray zone conflict and is constrained by the external deterrent threat, C will select

challenger responds to increases in the defender's willingness to go to war by scaling back their challenge to nothing, the challenger is clearly deterred. When the challenger responds to increases in the defender's willingness to go to war by engaging in more aggressive and riskier behavior, the defender's deterrent threat is not constraining the challenger, and we will say the challenger is choosing their challenge based on their internal efficiency.

Second, a similar dynamic can play out when the challenger begins by sometimes risking war. If the external deterrent threat constraint binds and κ decreases, C will either decreased their challenge (still selecting $g_C^* = g_C^{W,EDT}$, which is decreasing in κ), or revert to selecting into the status quo ($g_C^* = 0$). Alternatively, if the challenger sometimes risks war, the challenger's internal efficiency constraint binds, and κ decreases, then C will be non-responsive until κ decreases enough so that the external deterrent threat constraint binds.

How the challenger's responds to changes in κ can be summarized in Observations 1 and 2.

Observation 1: *If increasing the defender's willingness to go to war (decreasing κ) results in the challenger decreasing their selected challenge (g_C), then the external deterrent threat shapes the challenger's selected challenge.*

Observation 2: *If increasing the defender's willingness to go to war results in the challenger not changing their selected challenge, increasing their selected challenge, or risking war more, then the challenger's internal efficiency shapes the challenger's selected challenge.*

These Observations (and Observations 3-6 below) are not describing "tests" of the model. Rather, these Observations describe how, if we observe specific empirical patterns in the data, we can back out whether external deterrence or the challenger's internal optimization is driving the challenger's behavior. For example, Observation 1 suggests that if we observe a positive relationship between κ and g_C^* , then deterrence is playing a role in shaping the challenger's behavior. This Observation uses the comparative statics from the model to bridge what can be observed empirically to what motivates the scope of an actor's gray zone conflict levels.

$g_C^* = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$. As the defender becomes better at war, the challenger will either opt into the status quo and set $g_C^* = 0$, or select a greater challenge and risk war by setting $g_C^* = \min \left\{ \frac{(1-\omega)\theta}{2\beta_C}, \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta} \right\}$.

6.2 On the Challenger's Resolve and Conflict Intensity

When the internal efficiency constraint binds, the challenger selects their challenges based on how much they care about the asset and how costly it is to engage in more aggressive challenges. If we observe decreased challenges over issues that the challenger cares less about (approximated by a lower θ), then we know the challenger's decisions are based on their internal efficiency calculations. For example, suppose under one set of parameters the challenger will select a challenge that sometimes risks war and where the internal efficiency constraint binds ($g_C^* = g_C^{W,IE} = \frac{(1-\omega)\theta}{2\beta_C}$). As the challenger's resolve (θ) decreases, the challenger will limit their selected challenge in one of several different ways. First, the challenger may continue to risk war, but will select a more conservative challenge (i.e. still select $g_C^* = g_C^{W,IE}$, but this value is smaller). Alternatively, the challenger may no longer find it worthwhile to risk war and may switch into a more conservative challenge that always result in gray zone conflict ($g_C^* = g_C^{GZ,IE}$ or $g_C^* = g_C^{GZ,EDT}$). Finally, if the challenger's resolve decreases enough, the challenger no longer finds it worthwhile to challenge and will accept the status quo ($g_C^* = 0$). Together, when the challenger is responding to their internal efficiency, as the challenger's resolve decreases, the challenger will pursue less aggressive challenges.

If we observe less aggressive challenges over issues that challengers care more about, then we know the challenger's decisions are influenced by deterrence (and vice-versa). This (counter-intuitive) relationship stems from the bargaining component of the game. As the challenger's resolve increases, the defender must offer the challenger a greater policy concession in the fourth stage to avoid the challenger declaring war in the fifth stage.¹⁸ This greater concession in Stage 4 constrains how aggressive the challenge can be without triggering the defender to go to war in Stage 3. Specifically, whenever the deterrent threat binds, the challenger scales back their challenge, knowing they will be compensated later for doing so.¹⁹ For this reason, whenever the challenger is responding to the external deterrent threat, as the challenger's resolve decreases, the challenger will pursue more aggressive challenges.

We illustrate these dynamics in Figure 4. When C has a sufficiently low resolve (sufficiently low θ), C will not challenge and will accept the status quo. Moving to the right, as the challenger

¹⁸In equilibrium, $x = P(g_C, g_D) - \frac{\kappa_C}{\theta}$.

¹⁹This is formally captured in $g_C^{GZ,EDT} = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ and $g_C^{W,EDT} = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$

values the asset more, the challenger eventually begins engaging in challenges. Within the next equilibrium space (GZ, IE), the challenger's challenges are not aggressive enough to reach the defender's threshold for declaring war; this means that the internal efficiency constraint binds, and the challenge is increasing in the challenger's resolve. The selected challenge increases in θ until the challenger becomes constrained by the defender's willingness to go to war, or when the defender's deterrent threat binds (Gray Zone, EDT).²⁰ Within this region, as discussed above, the challenger's optimal challenge is decreasing in θ . Then eventually, the challenger's resolve increases to the cutpoint where the challenger becomes willing to risk war with the privately-good-at-war types of the defender in order to conduct a greater challenge against the privately-bad-at-war types. This is where war sometimes occurs, and there is a similar relationship between the challenger's internal efficiency (War, IE) and defender's deterrent threat (War, EDT).

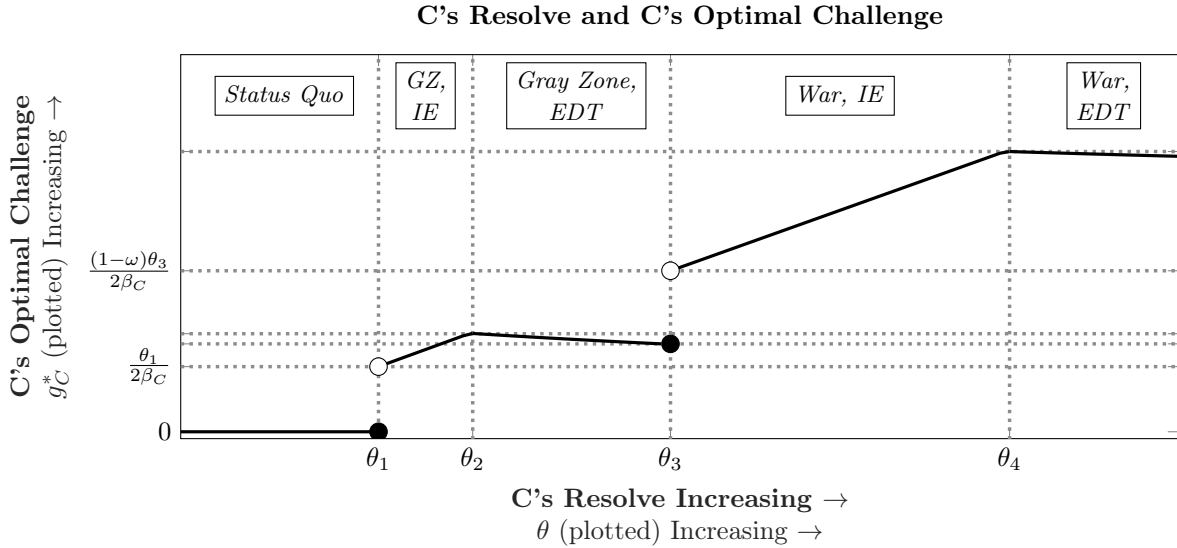


Figure 4: Optimal Challenge as C's Resolve Increases.

C's optimal challenge intensity under a range of θ 's. θ_1 is the value of θ where $\theta_1\rho - \kappa_C = \theta_1\rho + \frac{\theta_1^2}{4\beta_C} - \frac{\theta_1}{2\beta_D} - \kappa_C$, or C's gray zone payoffs are equal to C selecting into the status quo. θ_2 is where $\frac{\theta_2}{2\beta_C} = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta_2}$, or where, within gray zone conflict, C's internal efficiency constraint is equal to the D's deterrent threat constraint. θ_3 is where $\theta_3 \left(\rho - \frac{1}{4\beta_D} + \kappa_D \right) - \beta_C \left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta_3} \right)^2 = (1-\omega) \left(\theta_3\rho + \frac{(1-\omega)\theta_3^2}{2\beta_C} - \frac{\theta_3}{2\beta_D} - \kappa_C \right) + \omega (\theta_3\rho - \kappa_C) - \beta_C \left(\frac{(1-\omega)\theta_3}{2\beta_C} \right)^2$, or where C is indifferent between selecting into gray zone conflict and sometimes risking war. θ_4 is where $\frac{(1-\omega)\theta_4}{2\beta_C} = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta_4}$, or where, within cases where C sometime risks war, C's internal efficiency constraint is equal to the D's deterrent threat constraint.

²⁰This switch happens when $\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ becomes smaller than $\frac{\theta}{2\beta_C}$.

How the challenger's responds to changes in θ is summarized in Observations 3 and 4.

Observation 3: *If increasing the challenger's resolve (θ) results in the challenger increasing their selected challenge, then the challenger's internal efficiency shapes the challenger's selected challenge.*

Observation 4: *If increasing the challenger's resolve results in the challenger decreasing their selected challenge, then the external deterrent threat shapes the challenger's selected challenge.*

6.3 On the Challenger's Gray Zone Costs and Conflict Intensity

When the internal efficiency constraint binds, the challenger bases their challenges on how much they care about the asset and how costly it is to engage in more aggressive challenges. If we observe less aggressive challenges as the challenger's costs increases (approximated by greater β_C), then we know the challenger's decisions are based on their internal efficiency. In contrast, when the external deterrent threat constraint binds, the challenger bases their challenges on the defender's willingness to go to war, which is independent of C's internal costs to challenging. If we observe no change in the selected challenges as the challenger's costs are increasing, then we know the challenger's decisions are based on the external deterrent threat.

Across all equilibria, as the challenger's challenge costs decrease (β_C decreases), the challenger will select weakly more aggressive challenges (g_C^*). Figure 5 illustrates one example this. We do not discuss these results at length because these findings are relatively straightforward. How the challenger's responds to changes in β_C can be summarized in Observations 5 and 6.

Observation 5: *If decreasing the challenger's challenge costs (β_C) results in the challenger increasing their selected challenge, then the challenger's internal efficiency shapes the challenger's selected challenge.*

Observation 6: *If decreasing the challenger's challenge costs results in the challenger not changing their selected challenge, then the external deterrent threat shapes the challenger's selected challenge.*

There are two additional matters to note here. First, the results on β_C are quite different from the

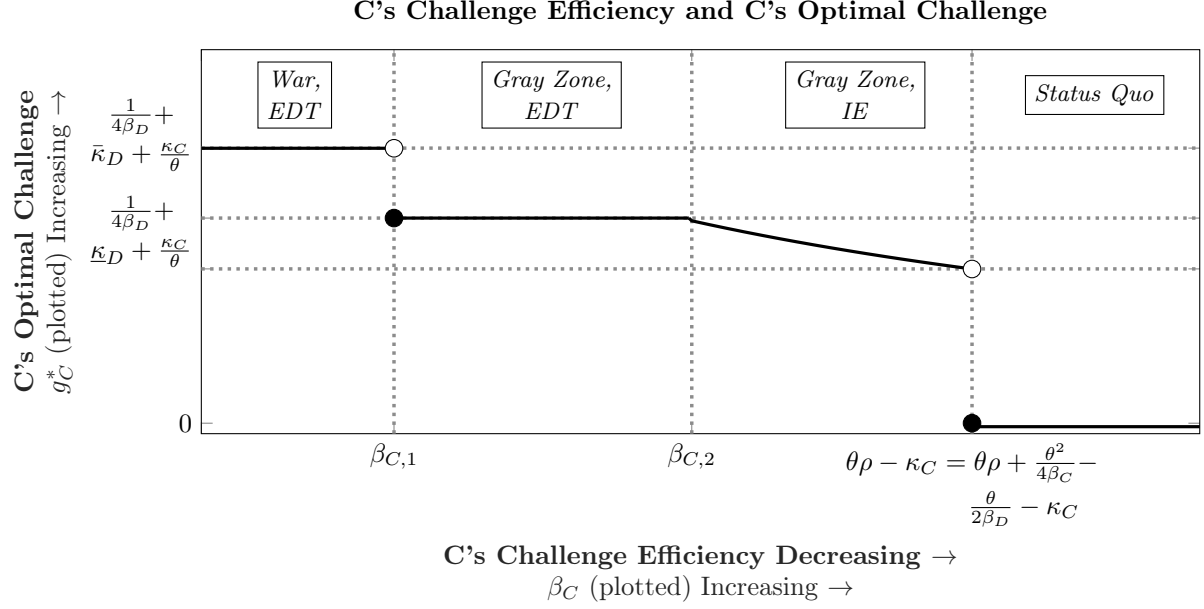


Figure 5: Optimal Challenge as C's costs of challenges vary.

C's selected challenge intensity under a range of β_C 's are plotted. $\beta_{C,1}$ is where $(1 - \omega)\theta \left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D \right) + \omega(\theta\rho - \kappa_C) - \beta_{C,1} \left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta} \right)^2 = \theta \left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D \right) - \beta_{C,1} \left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta} \right)^2$, or where C's gray zone payoffs are equal to C selecting into risking war sometimes. $\beta_{C,2}$ is where $\frac{\theta}{2\beta_{C,2}} = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$, or where the external deterrent threat constraint goes slack and the internal efficiency constraint begins to bind.

results on β_D , which is discussed below. This is the case because here, even if C goes to war, C faces costs to challenging (i.e. β_C is relevant) while D's payoffs are independent of β_D unless D goes to war (though this is relaxed in the extension in Section 2 the Appendix). Second, note that our model does not ever predict g_C^* to be increasing in β_C . This comparative static offers one way to “test” the model's findings. While we admit this is not a perfect test of the model, our empirical results are consistent with g_C^* decreasing in β_C .

7 Empirical Application: Russian Efforts in the Gray Zone

The empirical analysis assesses whether Russian gray zone behavior in Europe is shaped by NATO's deterrent threat or not. Observation 1 suggests that if we observe a negative relationship between NATO's willingness to fight and the intensity of Russian aggression, then deterrence is shaping Russian behavior. Below, our statistical analysis suggests that Russia appears to scale back its

challenges when the risk of escalation is high, which is consistent with the external deterrent threat binding at times (though this is subject to standard endogeneity concerns that exist in the multivariate regression setting). Additionally, we also find that Russian activity is inversely associated with decreased resolve and increased costs of its challenges, as approximated by a geographical “loss of strength gradient” (Posen 2003). Based on Observations 3 and 5, this suggests that Russian behavior is also driven by its internal efficiency constraint binding at times.

We focus on Russia because its interventions are referenced as paradigmatic examples of gray zone conflict (Marten 2015; Driscoll and Maliniak 2016; Jasper 2020). From 1994 to 2018, Russia has been involved in election interference in the United Kingdom and Moldova, cyberattacks in Estonia and Georgia, and special operations in Ukraine and Yugoslavia. The diversity of Russian targets and means provides an opportunity to conduct an analysis of Russian choices under different deterrent circumstances. We present a statistical analysis and brief qualitative analysis of three cases: Estonia, Ukraine, and Georgia. We find that the possibility of a NATO intervention is associated with more limited Russian gray zone operations. By no means is this analysis causal, as our key independent variables—NATO membership, for example—are not exogenous treatments. Our results should be viewed as suggestive and taken with the necessary caveats.

7.1 Data

Most quantitative studies of gray zone operations focus on a particular type, like cyber in the case of the Dyadic Cyber Incident and Dispute (DCID) data or electoral interference in the case of the Russian Electoral Interventions (REI) data (Valeriano and Maness 2014; Casey and Way 2017). Consequently, these data cover almost entirely distinct samples with significant differences concerning the severity of Russian attacks. To address these discrepancies, we construct a new, expanded dataset of 81 cases of Russian intervention from 1994 to 2018. DCID and REI together describe 72 unique cases of Russian aggression that have either included some degree of cyber intervention or were cases of electoral interference. We identify 9 additional instances of Russian cyberattacks in the coding period that are not listed in previous datasets. Most of these new cases involve cyber conflict after 2011 (the latest year in DCID) that were non-electoral (the universe of cases in REI). We also include the 3 cases of non-cyber Russian action from the International Crisis

Behavior (ICB) dataset (Brecher and Wilkenfeld 1997). For each incident, we create a new coding of the intensity of Russian attacks by coding whether Russia used five different types of military force in ascending order of intensity: (1) information operations (social media and disinformation), (2) cyber operations that result in disruption of infrastructure (service denial or industrial control system attacks), (3) overt use of special operations or unattributed military forces, (4) conventional air or sea forces, and (5) conventional ground forces. This data is then aggregated to the country-year level with a coding for the highest level of intensity of Russian intervention against each country in each year. Our dependent variable is thus an ordinal variable coded 5 for the highest level of intensity down to 0 for country-years experiencing no Russian attack.

Figure 6 plots the count and average intensity of Russian gray zone operations since 1994. There does not appear to be a clear temporal pattern in the intensity or frequency of activity; 2004 represents the most intense overall Russian interventions and 2014 experienced the highest number of interventions (most of which were associated with Ukraine).

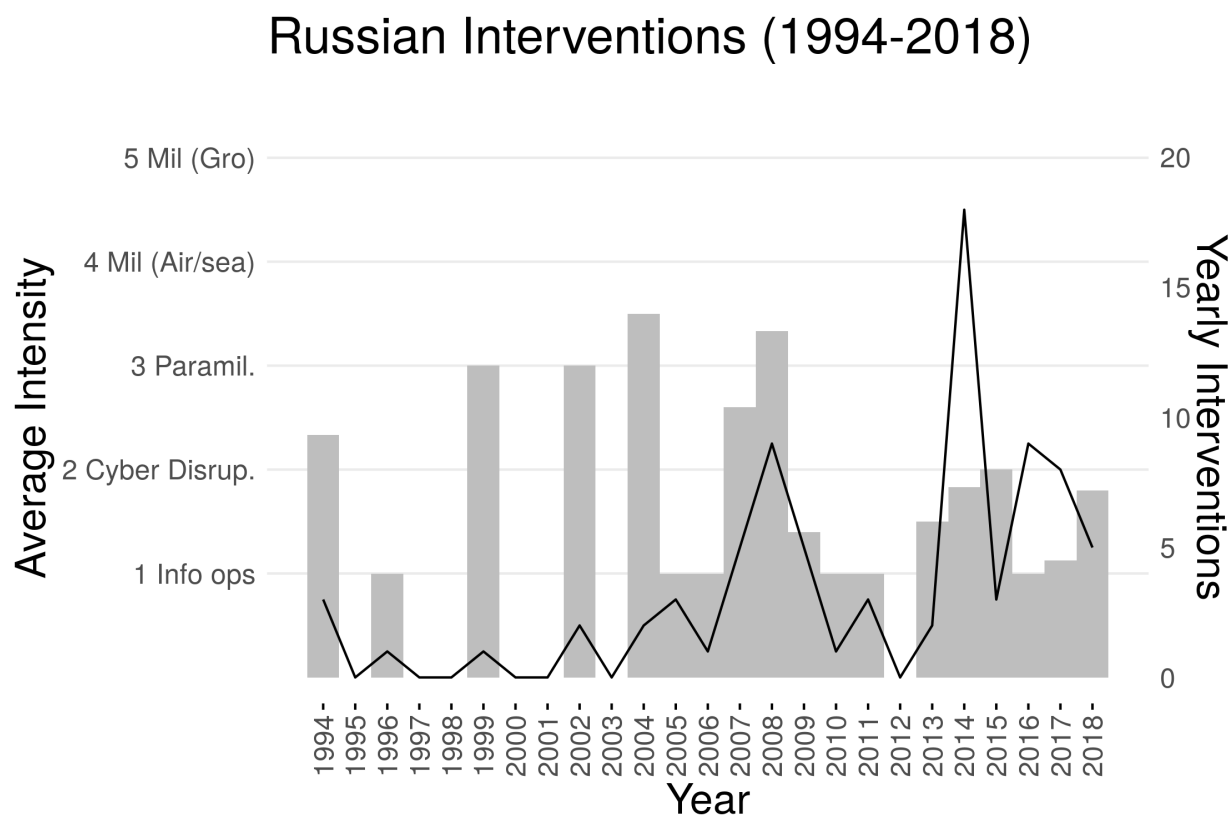


Figure 6: Intensity of Russian intervention across time. The line represents annual average intensity. Bars denote the number of interventions annually.

Figure 7 depicts a pattern of the geographical coverage of Russian conflict events in Europe. Russia appears to be willing to use more force in countries in its “near abroad,” relative to countries further away. Interpreting this geographic pattern on its own is difficult and highlights the need for more sophisticated analysis.

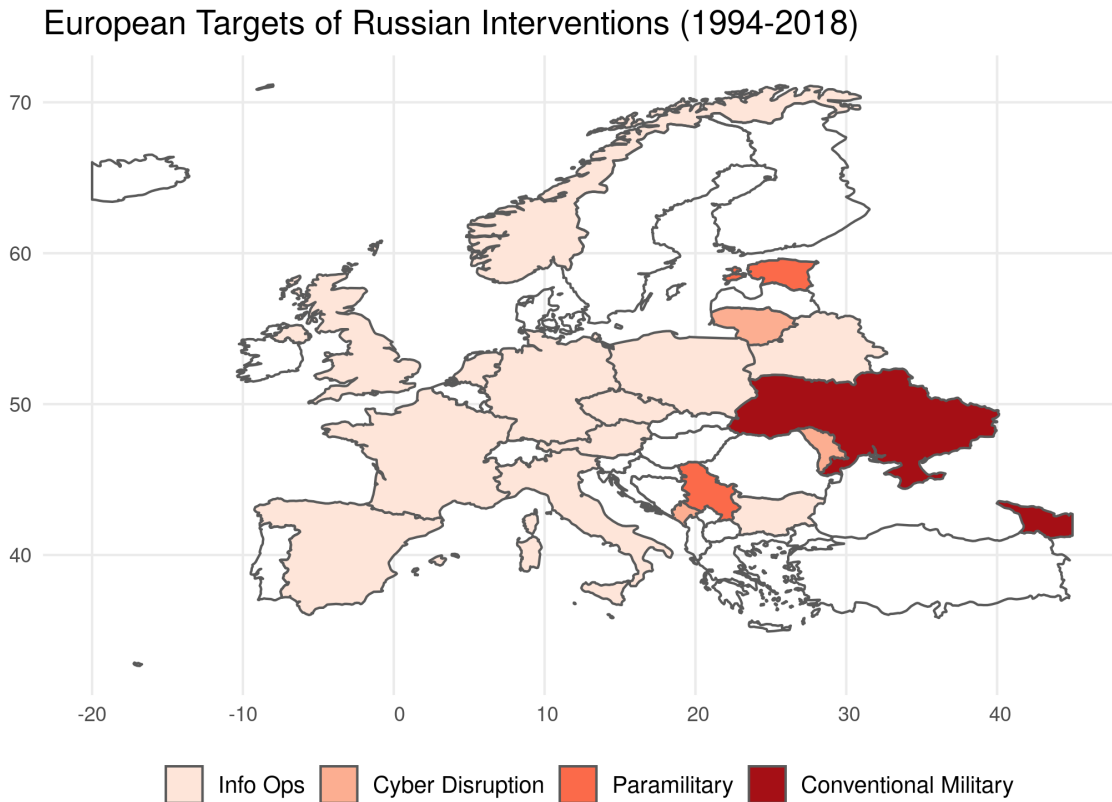


Figure 7: Geographic representation of Russia intervention. Shading represents the highest intensity of Russian intervention in each European state.

Now we discuss our independent variables. Consistent with the discussion on the external deterrent threat (Observation 1), we propose that NATO’s willingness (or unwillingness) to fight critically shapes Russian gray zone behavior. We operationalize this concept through a dummy variable for NATO membership. NATO members plausibly possess lower costs for fighting since they can rely on collective security, meaning they may possess a reduced willingness to tolerate aggressive low-level behavior from Russia. If Russia is responding to NATO’s deterrent threat, we expect NATO states to experience less intense Russian activity.

Consistent with the discussion of the internal-efficiency constraint in Observations 3 and 5, insofar

as military power is affected by a loss of strength gradient (Posen 2003), Russian gray zone operations will decrease as Russia has less resolve over the issue or faces greater costs for conducting operations when the internal efficiency constraint binds.²¹ We operationalize Russian resolve and gray zone efficacy jointly through a variable of the logged minimum distance between Russia and each potential-target state, as we expect Russia to care more about states on its periphery and to more easily conduct operations in its proximity (Weidmann, Kuse and Gleditsch 2010).

We also include a series of control variables motivated by existing theories of relationships between the control variables and our dependent and independent variables of interest. We include a democracy dummy for states with a Polity V score greater or equal to 6 to control for potential Russian eagerness to target democracies (Early and Asal 2018). A state’s possession of nuclear weapons may also alter Russia’s calculus about how to pursue aggressive actions, so we include a dummy variable for nuclear states (Gartzke and Kroenig 2009). We include GDP per capita and the log of population as larger, richer states could generate more opportunities for Russian interventions, especially cyber interventions (Beckley 2010). Finally, we include military expenditure since it influences both alliance decisions and the cost of undertaking aggression against an adversary (Omitoogun and Skons 2006). Summary statistics are included in the Appendix.

7.2 Empirical Model and Results

Because our outcome is an ordinal value, we estimate a series of ordered probit models with year fixed-effects and standard errors clustered by country. Our unit of analysis is the country-year. We run three empirical models on two samples. On the models, first we estimate the relationship between the intensity of Russian intervention and NATO membership and minimum distance from Russia without any control variables. Second, we re-run the first model while also controlling for the range of variables indicated above. We exclude military spending from the second model because it is missing across the entire panel for countries like Yugoslavia and Bosnia & Herzegovina. In model 3, we include the military spending variable, thus operationalizing military power using population and military spending. On the samples, our first sample (models (1)-(3)) includes all European states. We define state membership using the Gleditsch and Ward state list and continent

²¹We only analyze this variable in models that also include a NATO membership covariate.

location using the World Bank Development Indicator (Gleditsch and Ward 1999). This excludes micro-states with less than 250,000 people like Liechtenstein and San Marino. The downside is this sample may include states that are not of interest to Russia and may fall outside the scope of our model (i.e., Luxembourg). To address this, our second sample (models (4)-(6)) of “relevant European states,” includes only European states that meet any of the following three criteria: a) targets of a Russian attack from 1945-1993 as identified in the Militarized Interstate Dispute (MID) or International Crisis Behavior (ICB) datasets, b) Former Soviet Union or Warsaw Pact states, or c) states that are contiguous with Russia.

	Full sample			Relevant states sample		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Independent Variables						
NATO member	-0.28 (0.22)	-0.48** (0.20)	-0.62*** (0.22)	-0.47* (0.25)	-0.59** (0.26)	-0.69*** (0.25)
Russia distance	-0.10*** (0.04)	-0.11*** (0.03)	-0.12*** (0.03)	-0.05 (0.04)	-0.09** (0.04)	-0.09** (0.04)
Controls						
Democracy		0.15 (0.43)	0.45 (0.41)		0.11 (0.44)	0.43 (0.42)
Nuclear power		0.98** (0.45)	0.50 (0.47)		0.97* (0.52)	1.09 (0.84)
Population		0.18** (0.09)	0.13 (0.12)		0.15 (0.10)	0.18 (0.13)
GDP per cap		-0.01** (0.01)	-0.02** (0.01)		-0.01 (0.01)	-0.01 (0.01)
Mil. spending			0.02 (0.01)			-0.00 (0.02)
Observations	1,000	921	891	376	373	346

All models include year-fixed effects with country-clustered standard errors in parentheses. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table 2: Intensity of Russian Intervention: Ordered Probit Results

Table 2 presents the coefficient estimates from the ordered probit regressions run on both samples. The results show that both NATO membership and distance from Russia are associated with a decrease the intensity of Russian intervention against European states. Every model that utilizes control variables (models (2), (3), (5), and (6)) suggests the relationship is statistically significant at least at the 0.05 level.

These results provide evidence, subject to standard endogeneity concerns, that Russian gray zone behavior is shaped by NATO’s deterrent threat. Consistent with Observation 1, when facing NATO members, Russia undertook lower-intensity challenges. For example, model (3) reports a proportional odds ratio coefficient of 0.54 on the NATO dummy.²² This value means that for relevant NATO states, the odds of a non-cyber, non-information attack (categories 3, 4, or 5) are 46% lower than the odds of experiencing a cyber attack, an information attack, or no attack. Of course, our findings are also consistent with Russian behavior being shaped by their internal cost-benefit calculations (Observations 3 and 5). Naturally these findings can co-exist, suggesting that Russian motives for the scope of their gray zone conflict vary case-by-case. However, consistent with our earlier discussion, what we observe is consistent with deterrence plausibly playing a role.

Our findings are consistent across a range of alternate samples and modeling specifications detailed in the Appendix. As NATO membership is a somewhat coarse operationalization of external deterrent threat, we run additional models distinguishing between non-NATO members in the NATO accession membership process and those who are not.²³ Additional models also use distance from the nearest NATO state, as intervening in a state neighboring NATO may contain a higher risk of NATO response than interventions not bordering NATO. Additionally, we re-run the analysis including CINC ratios in place of population and the SIPRI military expenditure data (Singer, Bremer and Stuckey 1972).²⁴ The results from all these models are consistent with the original model specifications. We also find consistent results when sampling on only country-years that experience an attack, with OLS and ordered logit regression models, and using multiple imputation with additive regression, bootstrapping, and predictive mean matching to replace missing values for control variables thus mitigating concern that the initial results are an artifact of listwise deletion (van Buuren 2012; Arel-Bundock and Pelc 2018).

7.3 Case Studies

The quantitative analysis can be thought of as a coarse technique for considering empirical trends. For a more fine-grained analysis, we consider three major cyber campaigns attributed to Russia that

²²A complete table of all odds ratios is provided in the Appendix.

²³The pre-NATO stages includes membership in Partnership for Peace (PfP), Intensified Dialogue, and Membership Action Plan (MAP). For details on what these entail, see Amara and Paskevics (2010).

²⁴Population and military expenditure are two of the six components that comprise the CINC index.

feature prominently in the cybersecurity literature: Estonia, Georgia, and Ukraine. These cases are typically highlighted as examples of the increasing potential of gray zone conflict, making them a most likely case for the conventional efficiency logic. We follow a most similar case study design in selecting cases that feature cyber attacks by the same contiguous challenger (Russia) but differ in other military instruments employed (Bennett and Elman 2007).²⁵ Additionally, Russia has an interest in all states and could intervene with relative ease.

There are many potential explanations for Russian aspirations in each instance. Here we set aside Russia’s foreign policy formulation and focus instead on geopolitical context and military effectiveness. A summary of the extent of Russian intervention in each case is provided in Table 3; the columns are ordered geographically, West to East, rather than chronologically to highlight the deterrence gradient. Whereas our quantitative analyses proxy NATO membership and geographical distance for the deterrence gradient, our qualitative analysis allows us to add more nuance by considering relative Russian and Western interests.²⁶

Russian Operations	Estonia (2007)	Ukraine (2014)	Georgia (2008)
Conventional Forces			X
Special Operations		X	X
Cyber Operations	X	X	X

Table 3: Case comparison of Russian gray zone conflicts

7.3.1 Estonia (2007)

Estonia is the case with the most restrained Russian aggression. The Baltic states formally joined NATO in 2004 receiving only muted protests from Moscow. Later in 2007, as Putin’s foreign policy took on a more combative tenor, the relocation of a Soviet statue in Tallinn triggered a wave of DDoS attacks targeting Estonian banks and government sites (Schmidt 2013). A symbolic affront to Russian pride was met with symbolic protest. The attacks were carried out by “patriotic hackers” with tacit encouragement from Moscow and constituted one of the most disruptive cyber campaigns witnessed up to this time. Yet no one issued any clear demands or claimed responsibility, the attacks were effectively mitigated, and Estonia did not reinstall the statue.²⁷ Estonia’s defense minister

²⁵We include a fourth case study of Russian intervention in the 2016 US election in the appendix.

²⁶We discuss the complicating dynamics of Ukraine in 2022 in the next section.

²⁷In the context of our model, the relocated statue represented a new “status quo,” where Estonia could engage in independent, nationalist, possibly anti-Russian policy choices. The Russian cyberattack was a (largely futile) attempt

considered but ultimately rejected invoking Article V, the collective defense clause of the NATO treaty, instead treating the episode as a domestic law enforcement matter (Traynor 2007). Had Estonia not been a NATO member, this consideration would not even have been a possibility, and Russia might well have chosen to act more aggressively.²⁸ Although counterfactuals for the Estonian case are by definition not observable, the fact that Russia engaged in high intensity aggression against Moldova and Georgia during the 2005 to 2009 period while limiting its attacks against Lithuania, Poland, and the United States to cyber attacks and information operations provides suggestive evidence that Russia was clearly using much less force in cases where the defender was more capable or better protected than in cases where a target was weak and isolated politically. The unsettled legal status of a cyberattack in 2007 both enabled and constrained Russia in this respect (Joubert 2012). Russia and Estonia each understood that NATO was highly unlikely to retaliate for an ambiguous and plausibly deniable outburst, so long as Russia did not inflict serious harm and the level of aggression remained well below the perceived threshold for an Article V response. Overall, in the Estonian case Russian moves are consistent with a desire to avoid escalation. In the aftermath, Estonia overhauled its cyber defenses and established a NATO cybersecurity center, inhibiting future Russian provocations in the gray zone.

7.3.2 Georgia (2008)

Georgia is the case with the most Russian aggression. The NATO Bucharest Summit Declaration in April 2008 announced the potential for a Georgian and Ukrainian pathway to NATO membership. This, alongside Kosovo's declaration of independence, triggered a secession crisis in South Ossetia and Abkhazia. Russia announced that it would unilaterally increase peacekeepers in Abkhazia, and hostilities broke out later that year. Georgia was wracked by DDoS service attacks that were encouraged by Moscow, much like Estonia (Deibert, Rohozinski and Crete-Nishihata 2012). Unlike the Estonia case, Russia also deployed columns of conventional armor, naval bombardments, and special operations forces. This case is often cited as an example of how cyberspace can enhance military cross-domain operations, but even more important was Russia's relatively free hand. Russia chose whatever tools it needed to accomplish its objectives and did not pull its punches (Binnendijk

to undermine the Estonian government's ability to behave in this manner.

²⁸Estonians also may not have felt emboldened to provoke Russia in the first place without NATO protection.

2020). NATO counteraction in the Caucasus was unlikely because Georgia was deep in Russia's traditional sphere of influence and NATO members were publicly worried about Russia's reaction to Georgia joining NATO's Membership Action Plan (MAP). While Russia's tactical performance left much to be desired, the mission was a strategic success that ended the conversation about Georgia joining NATO. As Driscoll and Maliniak (2016, 590) point out, Georgia's location – with Tbilisi a mere 80 miles from the Russian border – makes it a security liability in the eyes of many Western leaders. Indeed, the Russian intervention served to clarify the stakes of Western interference in Russia's near abroad.

7.3.3 Ukraine (2014-2021)

The case of Ukraine, which has already been glossed above, falls in the middle. Consistent with the logic of NATO's deterrent threat binding, Russian actions in Ukraine (prior to 2022) were more extensive than those in Estonia, but much less than what occurred in Georgia. In eight years of protracted low-level conflict, there occurred neither large-scale combined arms warfare, as in Georgia, nor unrestrained ethnic cleansing (Driscoll and Steinert-Threlkeld 2020). Russia could have exerted greater military effort, but instead it emphasized cyber attacks, disinformation campaigns, special operations, and military aid to proxy forces. Even though NATO has no formal commitment to Ukraine, conflict in a country that borders four NATO allies (Poland, Slovakia, Hungary, Romania) is implicitly shaped by the possibility of Western intervention, risking nuclear escalation in the process. As a result, we conjecture that Russia acts circumspectly and relatively ineffectively. Endemic Russian cyberattacks and information operations have had little impact on battlefield events (Kostyuk and Zhukov 2019). Even as social media manipulation is supposedly a Russian specialty, pro-Kremlin narratives have not taken hold in Western Ukraine (Driscoll and Steinert-Threlkeld 2020). By and large, Russian subversion has failed to influence Kyiv (other than pushing it Westward) (Maschmeyer 2021).

7.3.4 Discussion of Cases

All three cases, each often cited to exemplify the potency of cyber-enabled gray zone conflict, are consistent with our hypothesis that deterrence encourages capable actors to engage in calculated restraint. As the deterrent threat from NATO becomes more salient moving westward from Georgia to

Ukraine and finally to Estonia, Russia pursues its international objectives with lower intensity.

One might argue that Russia has different levels of resolve across these cases. For example, one might argue that Russia places a very different value on the outcome in Ukraine than Estonia. Indeed, Russia let Estonia join NATO without a fight in 2004, and Russia’s historical ties with Estonia are weaker. By contrast, Russia had supported Georgian separatists since the early 1990s and was highly resolved to ward off Western encroachment. The comparison of Georgia and Ukraine, however, finds this alternative account wanting. While Georgia may be the birthplace of Stalin, it is also a peripheral outpost in the Caucasus far from Moscow. By contrast, the Black Sea port of Sevastopol makes Crimea strategically important, and NATO forces in Ukraine would be perceived as a dagger pointing at Moscow. The seat of the medieval Kievan Rus empire is also more salient in Russian nationalist mythology than Georgia ever was. If Russian moves were motivated by resolve rather than external deterrence, then we would expect more robust, and more overt, Russian military efforts in Ukraine. Yet, despite Russia’s undoubted higher valuation for the stakes in Ukraine, Russia showed more restraint there for eight years, at least until the February 2022 invasion (as we discuss in the next section).

8 Gray Zone Efficacy and Escalation Dynamics

Beyond microfounding the motives for gray zone conflict behavior, our model can also offer insight into an important policy question: how can a defender best react to the prospect of gray zone conflict?

It might seem intuitive that improvements in a defender’s capacity to counter gray zone aggression should reinforce the strength of deterrence, but this is not necessarily the case. As the defender becomes more capable at gray zone conflict (here approximated by lowering β_D), this constrains how productive gray zone conflict can be for the challenger. And, as the gray zone conflict option becomes worse for the challenger, the challenger will consider other policy options; sometimes the defender getting better at gray zone conflict will effectively deter the challenger—convince the challenger not to challenge and accept the “status quo”—but other times this change will encourage the challenger to behave more aggressively and risk a war. In terms of the risk-return trade-off

discussed in Section 5.1, consider a scenario where, in equilibrium, the challenger is choosing a conservative challenge that always results in gray zone conflict.²⁹ This means the challenger is not choosing a more aggressive challenge that would sometimes result in (a more productive form of) gray zone conflict and would sometimes result in war.³⁰ As the defender becomes better at gray zone conflict, gray zone conflict becomes less beneficial for the challenger; all the while, the challenger's war payoffs remain the same. This means that the “downside” of a more aggressive challenge—the game ending in war—does not look as bad to the challenger given that the challenger is now doing worse playing it safe.

This intuition can be visualized in Figure 8. The basic logic for why this figure looks as it does is described in Section 5.1, but it is worthwhile highlighting the value $\theta\rho - \kappa_C$ on the y-axis, which is C's expected payoff should the game end in war. In the top panel, when $\beta_D = 5$, C attains the greatest payoff setting $g_C^* = g_C^{GZ,EDT}$ and never risking war. It is worthwhile noting what C is not doing. C could risk war from setting $g_C = g_C^{W,EDT}$, which would result in a greater payoff if the game ends in gray zone conflict (when D is type $\bar{\kappa}_D$), but would also result in a much lower payoff if the game ends in war (when D is type $\underline{\kappa}_D$). Here, the downside of C setting $g_C = g_C^{W,EDT}$ and risking war—C having to fight a war—is so much worse than C's payoff of playing it safe and setting $g_C = g_C^{GZ,EDT}$, that C will not take the risk and will optimally set $g_C^* = g_C^{GZ,EDT}$.

In the bottom panel, when $\beta_D = 1.5$, C now does best setting $g_C = g_C^{W,EDT}$ and sometimes risking war. Why? As D becomes better at gray zone conflict, C does worse within all gray zone conflict outcomes. Visually, under $\beta_D = 1.5$, C setting $g_C = g_C^{GZ,EDT}$ is now much worse for C. So much worse, in fact, that the downside of C risking war—C's wartime payoff of $\theta\rho - \kappa_C$ —is only slightly less than C's payoff from setting $g_C = g_C^{GZ,EDT}$. Here, as β_D decreases, C does so poorly in gray zone conflict that now war does not look nearly as bad; this can incentivize C to tolerate the downside risk of war in order to attain a greater “upside” of a more aggressive gray zone challenge.

Note that the discussion above does not claim that decreases in the defender's gray zone costs always leads to war. As D gets better at gray zone conflict, C could instead be pressured out of challenging altogether and instead accept the status quo. A key factor that determines whether decreases in β_D

²⁹Formally, C is choosing $g_C^* = g_C^{GZ,EDT}$ or $g_C^* = g_C^{GZ,IE}$.

³⁰Formally, C is not choosing $g_C^* = g_C^{W,EDT}$ or $g_C^* = g_C^{W,IE}$.

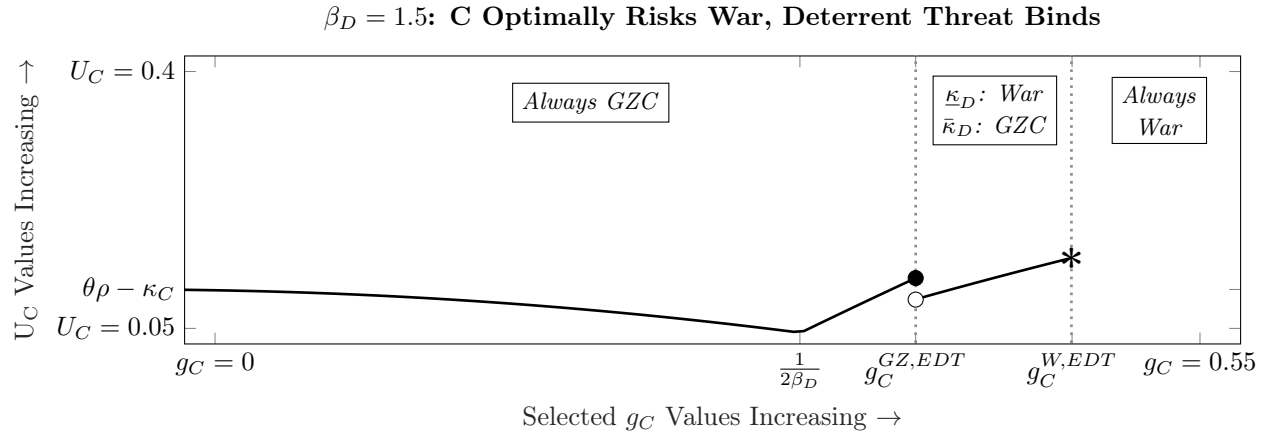
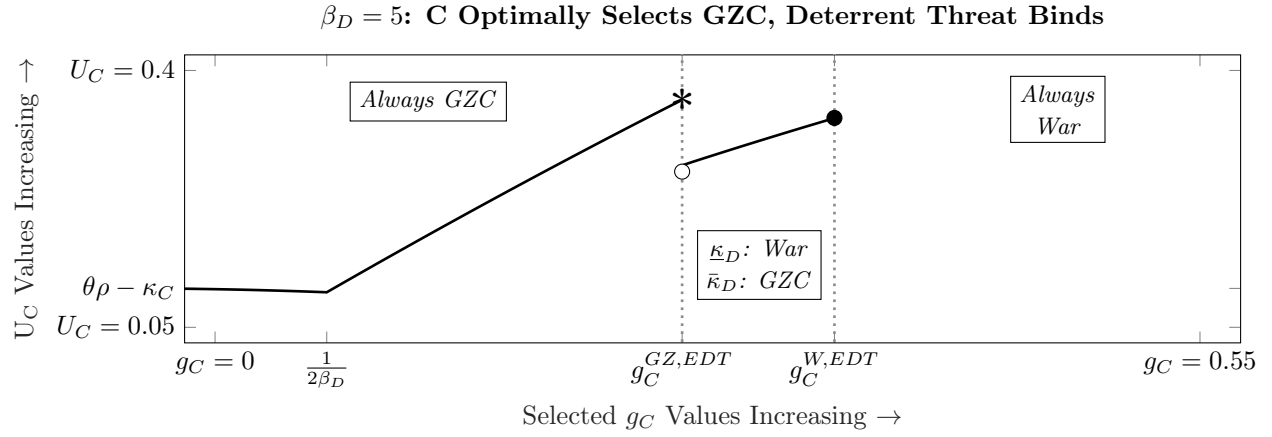


Figure 8: C's Utility Across selected g_C 's.

C's optimal g_C is marked by the asterisks. The regions denote how D responds to a selected g_C —whether D will always engage with a gray zone response, whether D will always go to war, or whether different types of D respond differently. All sub-figures have the same underlying parameters (listed in the Appendix), except β_C which varies across sub-figures.

result in C switching from gray zone conflict to war or peace is C's resolve. When the challenger possess a high resolve, a decrease in β_D will result in war under more circumstances. These finding can be summarized in Observation 7.

Observation 7: *Decreases in the defender's gray zone costs (β_D) can lead to war, especially when C's resolve (θ) is high.*

Until this point, we have discussed β_D as the defender's gray zone capabilities stemming from technological capabilities. There is an alternate interpretation as well; β_D could be influenced by the defender's prior, un-modeled moves in the game that set up current gray zone operations. For example, if the defender aggressively pursued counterinsurgency operations against foreign-backed rebels before this game began, the defender could be in a better position to pursue aggressive gray zone conflict within the game. This interpretation illustrates how it is not just latent, exogenous costs that influence the challenger's activity. Rather, war can result from the defender behaving too aggressively in prior gray actions against a highly resolved challenger.

This final interpretation has important policy consequences being played-out in Ukraine since February 2022. After roughly eight years of gray zone conflict in Ukraine, Russia dramatically changed course by invading Ukraine, thus initiating a highly aggressive war. This escalation raises an important question about gray zone conflict: why, if subversion is such an efficient way of revising the status quo without triggering open war, would Russia choose to accept the costs and risks of threatening just such a war? Our research poses one answer to this question. Over the past eight years, through internal improvements and external support, Ukraine has gotten better at fighting in the gray zone, which has made gray zone untenable for Russia. Our model suggests there could be a connection between improved Ukrainian gray zone capabilities (2014-2021) and the Russian escalation that followed. Of course, we cannot say for certain that this element drove Russia escalation; it is exceptionally difficult to identify what drives international decision making. For example, it could be that Russia came to view itself as more of a declining power in the lead up to the 2022 invasion, perhaps due to arms transfers to Ukraine or the warming relations between NATO and Ukraine (Benson and Smith 2022). However, even if we are incorrect in this one case, our model still suggests an important implication for future policy: challengers with high resolve, when faced

with worse options within the gray zone conflict policy space, may seek more aggressive means of conducting international politics.

There are three additional technical points worth mentioning. First, further discussion on this result, details on Observation 7, and comparative statics are found in the Appendix (Section 1.8). Second, reasonable readers might be concerned that the results are driven entirely by β_D not altering war payoffs. In the Appendix, we demonstrate a similar result to Observation 7 for a model where, as D gets better at gray zone conflict, D also gets better at war and C also gets worse at war. While the results are more complex and notation-heavy, the general mechanism remains (though with some limits). Third, a result like Observation 7 can also exist in a model where commitment problems (rather than incomplete information) are drivers of war. We explore this in Appendix Section 3.

9 Every Silver Lining’s Got a Touch of Gray

Gray zone conflict occurs when capable actors intentionally limit the intensity or capacity of aggression and refrain from escalation. Deterrence shapes the way that conflict emerges, but it may not suppress conflict altogether. The good news is that gray zone conflict may be symptomatic of deterrence success. Adversaries could be “designing around” deterrence in an effort to avoid the anticipated retaliation of the defender (Lieberman 2012). The bad news is that gray zone conflict probes the threshold of deterrence effectiveness. We expect conflict severity to be greater wherever there are questions about the willingness or ability of defenders to respond forcefully. An adversary is seldom passive. There will always be attempts at end-runs or push-back, even when deterrence is credible. It is thus important to think carefully before overextending commitments where credibility is in doubt.

Just as there is a gray zone between war and peace, the distinction between effective and ineffective deterrence is also fuzzy. Doubling down on deterrence in the gray zone can mitigate conflict in some cases but provoke escalation in the others. Wherever deterrence is credible, or the challenger’s resolve is low, challengers can be expected to exercise restraint as they probe to see what they can get away with. Wherever deterrence is not credible or challengers are highly resolved, however, a

challenger will be more emboldened to use whatever means they have at their disposal to meet their objectives, limited only by internal efficiency constraints.

We have used the same cases that have raised alarm about the dangers of gray zone conflict to suggest the validity of an alternative explanation. The evidence suggests that Russia may systematically reduce the intensity of its interventions due to deterrence, employing a greater variety of means with more lethal intensity where deterrence is weakest but conducting only ambiguous information operations where deterrence is most robust. The conventional wisdom is right that Russian interventions are paradigmatic exemplars of gray zone conflict, but it is wrong about Russian motivations and the effectiveness of these operations. Revisionist powers have not discovered a secret formula or novel tools destined to destabilize Western democracies or undermine NATO's deterrence posture. Rather it acts opportunistically as circumstances enable it to hassle adversaries and their clients without risking a costly military confrontation. The flip side of this logic, however, is that Russia is willing to call NATO's bluffs in cases where it can reasonably expect that NATO is unwilling to intervene. In Georgia, and even more so in Chechnya, Russian willingness to prioritize effectiveness at the price of efficiency is clear. Ukraine after 2021 provides a more nuanced, and more dangerous, illustration of this general logic.

The very fact that an adversary opts to engage in limited conflict suggests both vulnerabilities and opportunities. Instead of worrying about Western paralysis in response to Russian cunning, we can acknowledge that NATO has already blocked Russia from wielding even greater influence. NATO's implicit general deterrence posture has arguably succeeded in keeping more extreme forms of Russian aggression in check. The unfortunate fact remains, however, that a simple remedy for gray zone conflict does not exist, if only because there is always some ability to "design around" deterrence. We may be able to choose *how* our adversary confronts us, but not *whether*. Deterrence is not an on/off switch, but a rheostat, providing a range of causes and variable effects. Gray zone conflict is, then, not simply deterrence failure but also a modest kind of success.

References

- Altman, Dan. 2018. "Advancing without Attacking: The Strategic Game around the Use of Force." *Security Studies* 27(1):58–88.
- Amara, Jomana and Martins Paskevics. 2010. "Unfulfilled Promises: The Impact of Accession on Military Expenditure Trends for New NATO Members." *Comparative Strategy* 29(5):432–449.
- Arel-Bundock, Vincent and Krzysztof J. Pelc. 2018. "When Can Multiple Imputation Improve Regression Estimates?" *Political Analysis* 26(2):240–245.
- Baliga, Sandeep, Ethan Bueno De Mesquita and Alexander Wolitzky. 2020. "Deterrence with Imperfect Attribution." *American Political Science Review* 114(4):1155–1178.
- Beckley, Michael. 2010. "Economic Development and Military Effectiveness." *Journal of Strategic Studies* 33(1):43–79.
- Bennett, Andrew and Colin Elman. 2007. "Case Study Methods in the International Relations Subfield." *Comparative Political Studies* 40(2):170–195.
- Benson, Brett V. and Bradley C. Smith. 2022. "Commitment Problems in Alliance Formation." *American Journal of Political Science* n/a(n/a).
- Binnendijk, Anika. 2020. Understanding Russian Black Sea Power Dynamics Through National Security Gaming. Technical report RAND Corporation Santa Monica, CA: .
- Bragg, Belinda. 2017. Integration Report: Gray Zone Conflicts, Challenges, and Opportunities. Technical report Arlington, VA: .
- Brands, Hal. 2016. Paradoxes of the Gray Zone. SSRN Scholarly Paper ID 2737593 Social Science Research Network Rochester, NY: .
- Brecher, Michael and Jonathan Wilkenfeld. 1997. *A Study of Crisis*. University of Michigan Press.
- Brodie, Bernard. 1957. "More About Limited War." *World Politics* 10(1):112–122.
- Carcelli, Shannon and Erik A. Gartzke. 2017. The Diversification of Deterrence: New Data and Novel Realities. In *Oxford Research Encyclopedia of Politics*. Oxford University Press.

- Carson, Austin. 2016. "Facing Off and Saving Face: Covert Intervention and Escalation Management in the Korean War." *International Organization* 70(1):103–131.
- Carson, Austin. 2018. *Secret Wars: Covert Conflict in International Politics*. Princeton Studies in International History and Politics Princeton, NJ: Princeton University Press.
- Casey, Adam and Lucan Ahmad Way. 2017. "Russian Electoral Interventions, 1991-2017."
- Clarke, Kevin A. and David M. Primo. 2012. *A Model Discipline: Political Science and the Logic of Representations*. OUP USA.
- Coe, Andrew J. 2018. "Containing Rogues: A Theory of Asymmetric Arming." *The Journal of Politics* 80(4):1197–1210.
- Danilovic, Vesna and Joe Clare. 2010. Deterrence and Crisis Bargaining. In *Oxford Research Encyclopedia of International Studies*. Oxford University Press.
- Debs, Alexandre and Nuno P. Monteiro. 2014. "Known Unknowns: Power Shifts, Uncertainty, and War." *International Organization* 68(1):1–31.
- Deibert, Ronald, Rafal Rohozinski and Masashi Crete-Nishihata. 2012. "Cyclones in Cyberspace: Information Shaping and Denial in the 2008 Russia–Georgia War." *Security Dialogue* 43(1):3–24.
- Driscoll, Jesse and Daniel Maliniak. 2016. "With Friends Like These: Brinkmanship and Chain-Ganging in Russia's Near Abroad." *Security Studies* 25(4):585–607.
- Driscoll, Jesse and Zachary Steinert-Threlkeld. 2020. "Social Media and Russian Territorial Irredentism: Some Facts and a Conjecture." *Post-Soviet Affairs* 36(2):101–121.
- Dunford, Joseph. 2016. "Gen. Dunford's Remarks and Q&A."
- Early, Bryan and Victor Asal. 2018. "Nuclear Weapons, Existential Threats, and the Stability–Instability Paradox." *The Nonproliferation Review* 25(3-4):223–247.
- Fallon, Michael. 2017. "Speech Delivered by Secretary of State for Defence Sir Michael Fallon at the RUSI Landwarfare Conference."

- Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.
- Fearon, James D. 1997. "Signaling Foreign Policy Interests Tying Hands versus Sinking Costs." *Journal of Conflict Resolution* 41(1):68–90.
- Freedman, Lawrence. 2004. *Deterrence*. United Kingdom: Wiley.
- Gannon, J Andrés. 2022. "One If by Land, and Two If by Sea: Cross-domain Contests and the Escalation of International Crises." *International Studies Quarterly* 66(4).
- Gartzke, Erik and Matthew Kroenig. 2009. "A Strategic Approach to Nuclear Proliferation." *Journal of Conflict Resolution* 53(2):151–160.
- George, Alexander and Richard Smoke. 1989. "Deterrence and Foreign Policy." *World Politics* 41(2):170–182.
- Gleditsch, Kristian S. and Michael D. Ward. 1999. "A Revised List of Independent States since the Congress of Vienna." *International Interactions* 25(4):393–413.
- Goldstein, Lyle. 2017. "The US–China Naval Balance in the Asia-Pacific: An Overview." *The China Quarterly* 232:904–931.
- Green, Michael, Kathleen Hicks, Zack Cooper, John Schaus and Jake Douglas. 2017. *Countering Coercion in Maritime Asia: The Theory and Practice of Gray Zone Deterrence*. Rowman & Littlefield.
- Gurantz, Ron and Alexander V. Hirsch. 2017. "Fear, Appeasement, and the Effectiveness of Deterrence." *The Journal of Politics* 79(3):1041–1056.
- Hart, Sir Basil Henry Liddell. 1967. *Strategy: The Indirect Approach*. London: Faber & Faber.
- Hicks, Kathleen H. and Alice Hunt Friend. 2019. By Other Means Part I: Campaigning in the Gray Zone. Technical report Center for Strategic & International Studies Lanham: .
- Hoffman, Frank G. 2007. Conflict in the 21st Century: The Rise of Hybrid Wars. Technical report Potomac Institute for Policy Studies Arlington, VA: .

- Holmes, James R. and Toshi Yoshihara. 2017. "Deterring China in the "Gray Zone": Lessons of the South China Sea for U.S. Alliances." *Orbis* 61(3):322–339.
- Hoover, Kevin D. 2006. The Methodology of Econometrics. In *Palgrave Handbook of Econometrics. 1: Econometric Theory*, ed. Terence C. Mills and Kerry Patterson. First ed. Basingstoke: Palgrave Macmillan pp. 61–87.
- Hughes, Geraint. 2020. "War in the Grey Zone: Historical Reflections and Contemporary Implications." *Survival* 62(3):131–158.
- Huth, Paul K. 1999. "Deterrence and International Conflict: Empirical Findings and Theoretical Debates." *Annual Review of Political Science* 2(1):25–48.
- Jackson, Colin F. 2016. "Information Is Not a Weapons System." *Journal of Strategic Studies* 39(5-6):820–846.
- Jackson, Van. 2017. "Tactics of Strategic Competition: Gray Zones, Redlines, and Conflict before War." *Naval War College Review* 70(3):39–61.
- Janičatová, Silvie and Petra Mlejnková. 2021. "The Ambiguity of Hybrid Warfare: A Qualitative Content Analysis of the United Kingdom's Political–Military Discourse on Russia's Hostile Activities." *Contemporary Security Policy* .
- Jasper, Scott. 2020. *Russian Cyber Operations: Coding the Boundaries of Conflict*. Georgetown University Press.
- Jervis, Robert. 1984. *The Illogic of American Nuclear Strategy*. Cornell University Press.
- Joubert, Vincent. 2012. Five Years after Estonia's Cyber Attacks: Lessons Learned for NATO? Technical Report 76 NATO Defense College Rome, Italy: .
- Kenkel, Brenton and Peter Schram. 2021. "Uncertainty in Crisis Bargaining with Multiple Policy Options: A Game-Free Analysis."
- Kissinger, Henry. 1955. "Military Policy and Defense of the "Grey Areas"." *Foreign Affairs* 33(3):416–428.

- Kissinger, Henry. 1957. "Strategy and Organization." *Foreign Affairs* 35(3):379–394.
- Kostyuk, Nadiya and Yuri Zhukov. 2019. "Invisible Digital Front: Can Cyber Attacks Shape Battlefield Events?" *Journal of Conflict Resolution* 63(2):317–347.
- Lanoszka, Alexander. 2016. "Russian Hybrid Warfare and Extended Deterrence in Eastern Europe." *International Affairs* 92(1):175–195.
- Lieberman, Elli. 2012. *Reconceptualizing Deterrence: Nudging Toward Rationality in Middle Eastern Rivalries*. Routledge.
- Lindsay, Jon R. and Erik A. Gartzke. 2018. Coercion through Cyberspace: The Stability-Instability Paradox Revisited. In *Coercion: The Power to Hurt in International Politics*, ed. Kelly M. Greenhill and Peter Krause. New York, NY: Oxford University Press.
- Marten, Kimberly. 2015. "Putin's Choices: Explaining Russian Foreign Policy and Intervention in Ukraine." *The Washington Quarterly* 38(2):189–204.
- Maschmeyer, Lennart. 2021. "The Subversive Trilemma: Why Cyber Operations Fall Short of Expectations." *International Security* 46(2):51–90.
- Matissek, Jahara W. 2017. "Shades of Gray Deterrence: Issues of Fighting in the Gray Zone." *Journal of Strategic Security* 10(3):1–26.
- Mazarr, Michael. 2015. "Struggle in the Gray Zone and World Order." <http://warontherocks.com/2015/12/struggle-in-the-gray-zone-and-world-order/>.
- McCarthy, Michael C., Matthew A. Moyer and Brett H. Venable. 2019. Deterring Russia in the Gray Zone. Technical report Strategic Studies Institute Carlisle Barracks, PA: .
- McCormack, Daniel and Henry Pascoe. 2017. "Sanctions and Preventive War." *Journal of Conflict Resolution* 61(8):1711–1739.
- Olson, Erik. 2015. "America's Not Ready for Today's Gray Wars." <https://www.defenseone.com/ideas/2015/12/americas-not-ready-todays-gray-wars/124381/>.

- Omitoogun, Wuyi and Elisabeth Skons. 2006. Military Expenditure Data: A 40-Year Overview. In *SIPRI Yearbook 2006: Armaments, Disarmament and International Security*. Stockholm International Peace Research Institute pp. 269–294.
- O’Rourke, Lindsey. 2018. *Covert Regime Change: America’s Secret Cold War*. Cornell Studies in Security Affairs Ithaca, NY: Cornell University Press.
- Osgood, Robert. 1969. “The Reappraisal of Limited War.” *The Adelphi Papers* 9(54):41–54.
- Pettyjohn, Stacie L. and Becca Wasser. 2019. Competing in the Gray Zone: Russian Tactics and Western Responses. Technical report RAND Corporation Santa Monica, CA: .
- Posen, Barry R. 2003. “Command of the Commons: The Military Foundation of U.S. Hegemony.” *International Security* 28(1):5–46.
- Powell, Robert. 2015. “Nuclear Brinkmanship, Limited War, and Military Power.” *International Organization* 69(3):589–626.
- Quackenbush, Stephen L. 2011. “Deterrence Theory: Where Do We Stand?” *Review of International Studies* 37(2):741–762.
- Ramsay, Kristopher W. 2017. “Information, Uncertainty, and War.” *Annual Review of Political Science* 20(1):505–527.
- Schelling, Thomas. 1966. *Arms and Influence*. Yale University Press.
- Schmidt, Andreas. 2013. The Estonian Cyberattacks. In *A Fierce Domain: Conflict in Cyberspace, 1986 to 2012*, ed. Jason Healey. Cyber Conflict Studies Association pp. 174–193.
- Schram, Peter. 2021. “Hassling: How States Prevent a Preventive War.” *American Journal of Political Science* 62(2):294–308.
- Schultz, Kenneth A. 2010. “The Enforcement Problem in Coercive Bargaining: Interstate Conflict over Rebel Support in Civil Wars.” *International Organization* 64(2):281–312.
- Shultz, George. 1986. “Low-Intensity Warfare: The Challenge of Ambiguity.”

- Singer, David, Stuart Bremer and John Stuckey. 1972. *Capability Distribution, Uncertainty, and Major Power War, 1820-1965*. Sage Publications pp. 19–48.
- Slantchev, Branislav L. 2011. *Military Threats: The Costs of Coercion and the Price of Peace*. Cambridge University Press.
- Snyder, Glenn H. 1965. The Balance of Power and the Balance of Terror. In *World in Crisis: Readings in International Relations*, ed. Frederick Hartmann. New York: The Macmillan Company pp. 180–191.
- Spaniel, William. 2019. *Bargaining over the Bomb: The Successes and Failures of Nuclear Negotiations*. Cambridge University Press.
- Tarar, Ahmer. 2016. “A Strategic Logic of the Military Fait Accompli.” *International Studies Quarterly* 60(4):742–752.
- Thornton, Rod. 2015. “The Changing Nature of Modern Warfare.” *The RUSI Journal* 160(4):40–48.
- Tor, Uri. 2015. “‘Cumulative Deterrence’ as a New Paradigm for Cyber Deterrence.” *Journal of Strategic Studies* 40(1-2):92–117.
- Traynor, Ian. 2007. “Russia Accused of Unleashing Cyberwar to Disable Estonia.” *The Guardian* .
- Turbiville, Graham. 2002. “Preface: Future Trends in Low Intensity Conflict.” *Low Intensity Conflict & Law Enforcement* 11(2-3):155–163.
- Valeriano, Brandon and Ryan C Maness. 2014. “The Dynamics of Cyber Conflict between Rival Antagonists, 2001–11.” *Journal of Peace Research* 51(3):347–360.
- van Buuren, Stef. 2012. *Flexible Imputation of Missing Data*. CRC Press.
- Weidmann, Nils, Doreen Kuse and Kristian Skrede Gleditsch. 2010. “The Geography of the International System: The CShapes Dataset.” *International Interactions* 36(1):86–106.
- Werner, Suzanne. 2000. “Deterring Intervention: The Stakes of War and Third-Party Involvement.” *American Journal of Political Science* 44(4):720–732.

- Wirtz, James J. 2017. "Life in the "Gray Zone": Observations for Contemporary Strategists." *Defense & Security Analysis* 33(2):106–114.
- Zagare, Frank C. 1996. "Classical Deterrence Theory: A Critical Assessment." *International Interactions* 21(4):365–387.
- Zagare, Frank C. and D. Marc Kilgour. 1998. "Deterrence Theory and the Spiral Model Revisited." *Journal of Theoretical Politics* 10(1):59–87.